

U.S. DEPARTMENT OF COMMERCE
Office of Inspector General



***NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION***

*Improvements Needed in the Reporting of
Performance Measures Related to Goals for
Advancing Short-Term Warnings and
Implementing Seasonal to Interannual
Climate Forecasts*

Audit Report No. FSD-15643-3-0001/September 2003

PUBLIC RELEASE

Office of Audits, Financial Statements Audits Division



UNITED STATES DEPARTMENT OF COMMERCE
The Inspector General
Washington, D.C. 20230

SEP 30 2003

MEMORANDUM FOR: Vice Admiral Conrad C. Lautenbacher, Jr., USN (Ret.)
Under Secretary of Commerce for Oceans and Atmosphere
National Oceanic and Atmospheric Administration

FROM:

Johnnie E. Frazier

SUBJECT:

*Improvements Needed in the Reporting of Performance
Measures Related to Goals for Advancing Short-Term
Warnings and Implementing Seasonal to Interannual
Climate Forecasts*

Final Audit Report No. FSD-15643-3-0001

This is our final report on select performance measures used by NOAA to support two of its performance goals: (1) advance short-term warnings and forecasts, and (2) implement seasonal to interannual climate forecasts. The goals largely reflect activities of the National Weather Service (NWS), National Environmental Satellite, Data, and Information Service (NESDIS), and Office of Oceanic and Atmospheric Research (OAR) and are included in the Department of Commerce's *FY 2001 Annual Program Performance Report / FY 2003 Annual Performance Plan, FY 2001 Accountability Report, and FY 2002 Performance & Accountability Report*.

We found that NOAA is committed to reporting outcome-oriented measures and reliable information, including the use of extensive verification procedures by NWS for its performance measures relating to severe weather warnings. However, we identified instances in which reported information was at times incomplete, inaccurate, or unclear—often the result of inadequate explanations of and disclosures for the measures, coupled with some lapses in internal controls. Rectifying these problems would improve the usefulness and reliability of this information for Congress and OMB, both of whom rely on this data as part of the budget process.

In responding to the draft report, NOAA either concurred with or is taking corrective action consistent with all of the recommendations. For each recommendation, NOAA identified corrective actions taken or planned and implementation schedules. However, NOAA expressed concern that (1) the tone and tenor of the draft report was overly negative relative to the findings and recommendations presented and (2) the sample size used to examine performance data for tornado and flash flood lead times was not large enough to characterize identified deficiencies as internal control weaknesses. Nevertheless, the NWS response includes planned actions to strengthen internal controls

and reduce the likelihood of the kinds of lapses in internal controls identified during our review. NOAA also suggested that we include additional discussion of the process used by the agency to ensure the accuracy of performance information in the report.

Where appropriate, we have modified the report to reflect NOAA's response. However, we believe the tone and tenor of the final report is consistent with the need for NOAA to improve the completeness, accuracy, and clarity of performance information it reports. Within the appropriate sections of this report, we summarize NOAA's response to our draft report as well as provide our comments. NOAA's complete response is attached to the report as Appendix I.

In accordance with Department Administrative Order 213-5, please provide us with your action plan addressing the recommendations for our review and concurrence within 60 days of this memorandum. Should you need to discuss the content of this report and the action plan, please contact me at (202) 482-4661; or Michael Sears, Assistant Inspector General for Auditing, at (202) 482-1934.

We appreciate the cooperation and courtesies your staff extended to us during our review.

Attachment

cc: Jack J. Kelly, Jr.
Deputy Undersecretary of Commerce for
Oceans and Atmosphere

Gregory W. Withee
Assistant Administrator for
Satellite and Information Services

Louisa Koch
Acting Assistant Administrator
Office of Oceanic and Atmospheric Research

William F. Broglie
Chief Administrative Officer

Ted David
Acting Chief Financial Officer

Mack Cato
Director
Audits, Internal Control and Information Management

TABLE OF CONTENTS

	Page
EXECUTIVE SUMMARY	i
INTRODUCTION	1
OBJECTIVES, SCOPE, AND METHODOLOGY	4
FINDINGS AND RECOMMENDATIONS	6
I. Performance Measure: Lead Time (Minutes), Accuracy (%), and False Alarm Rate (FAR%) for Severe Weather Warnings for Tornadoes	7
A. Average lead time does not convey the full picture of NWS performance	7
B. Internal controls over generating and reporting data should be strengthened	8
C. Consistent procedures for maintaining supporting documentation need to be established and enforced	9
D. Additional disclosures would enhance the usefulness of reported results	9
E. Not implementing certain quality control procedures could have impacted reliability	9
F. Recommendations	10
G. NOAA Response	10
H. OIG Comments	11
II. Performance Measure: Lead Time (Minutes) and Accuracy (%) for severe weather warnings for flash floods	12
A. Average lead time does not fully capture performance	12
B. Internal controls should be strengthened	13

C.	Consistent procedures for maintaining supporting documentation need to be established and enforced	14
D.	Additional disclosures would enhance the usefulness of reported results	14
E.	Not implementing certain quality control procedures could have impacted reliability	15
F.	Recommendations	15
G.	NOAA Response	15
H.	OIG Comments	16
III.	Performance Measure: Accuracy (%) of Three-day Forecast of Precipitation	17
A.	Improvements are needed to more accurately measure the forecasting performance	18
B.	Additional disclosures are needed to improve presentation	18
C.	Recommendations	19
D.	NOAA Response	19
E.	OIG Comments	19
IV.	Performance Measure: Determine the Accuracy of the Correlation Between Forecasts of the Southern Oscillation Index (SOI) and El Niño/La Niña Events	20
A.	Usefulness of measure is hampered by limited, inadequate discussion of results	20
B.	Recommendations	21
C.	NOAA Response	21
D.	OIG Comments	21

V.	Performance Measure: U.S. Temperature Forecasts (Skill Score)	22
A.	Performance results were calculated using different time spans	22
B.	Additional disclosures are needed to improve presentation	23
C.	Recommendations	23
D.	NOAA Response	24
E.	OIG Comments	24
VI.	Performance Measure: New Monitoring or Forecast Products that Become Operational Per Year (cumulative)	25
A.	NOAA needs a consistent definition for a monitoring or forecast product	25
B.	NOAA has no verification procedures for new products	25
C.	Recommendations	26
D.	NOAA Response	26
E.	OIG Comments	26
VII.	Performance Measure: New Climate Observations Introduced	27
A.	Reported results are innacurate	27
B.	Measure does not convey what is being reported	28
C.	Recommendations	28
D.	NOAA Response	29
E.	OIG Comments	29

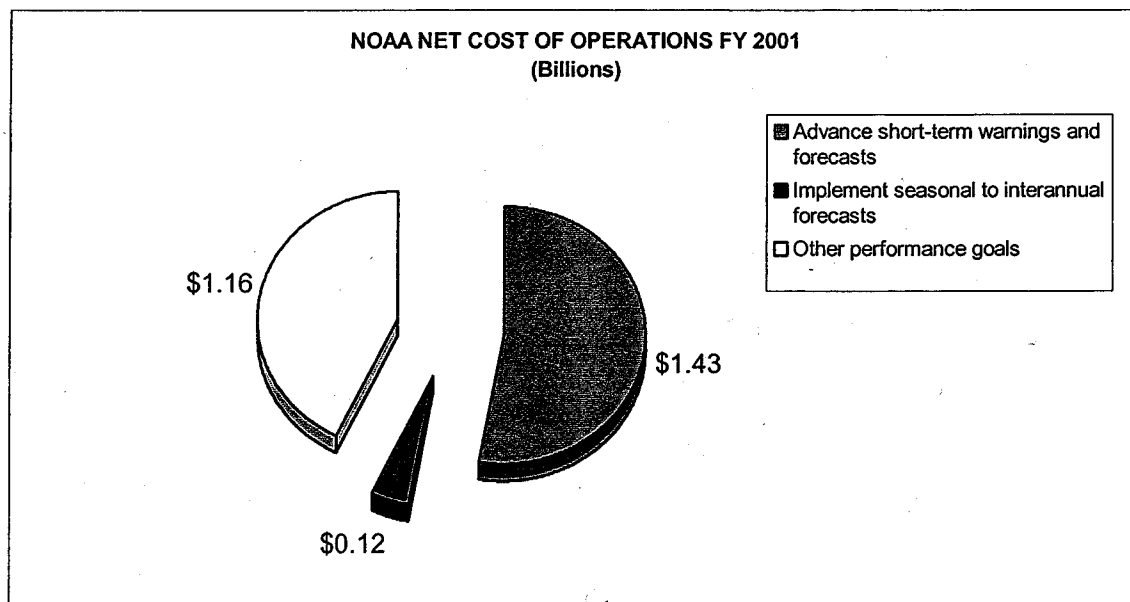
APPENDIX I

EXECUTIVE SUMMARY

The National Oceanic and Atmospheric Administration (NOAA) is charged with assessing and predicting changes in the Earth's environment, and protecting and managing marine and coastal resources to ensure sustainable economic opportunities—missions that support the Department's strategic goal of observing and managing the Earth's environment to promote sustainable growth.¹ As such, NOAA's performance plans, program results, and financial information are integral components of Commerce's annual performance plans and reports submitted to meet the requirements of the Government Performance and Results Act (GPRA) of 1993.

NOAA maintains seven performance goals to support the Department's strategic goal: (1) build sustainable fisheries, (2) sustain healthy coasts, (3) recover protected species, (4) advance short-term warnings and forecasts, (5) implement seasonal to interannual climate forecasts, (6) predict and assess decadal to centennial change, and (7) promote safe navigation. It has established a number of measures to gauge its success at achieving each goal.

We conducted a performance audit of select measures that support two of these goals—advance short-term warnings and forecasts, and implement seasonal to interannual climate forecasts—to (1) assess the collection and reporting of NOAA performance information in documents submitted to meet GPRA requirements, and (2) determine whether NOAA's internal controls are sufficient to ensure that performance data is accurate, consistent, and reliable. Net cost of operations for the two goals for the year ended FY 2001 was more than \$1.5 billion.²



¹ This is one of three Department of Commerce strategic goals. The other two are (1) provide the information and framework to enable the economy to operate efficiently and equitably, and (2) provide infrastructure for innovation to enhance American competitiveness.

Over the past several years, NWS has received accolades for its collection and reporting of performance results. Similarly, we noted a commitment on the part of NOAA to report outcome-oriented measures and reliable information, including extensive efforts by the NWS to ensure the accuracy of performance information relating to its severe weather warnings. However, our review found that (1) reported performance data at times did not provide a complete picture of performance; (2) performance information was not always accurately reported, and (3) explanatory language frequently did not appropriately describe results or the limitations of the data. Improvements were needed for each of the seven measures we reviewed, as follows:

Performance Goal: Advance short-term warnings and forecasts

1. **Measure: Lead time (minutes), accuracy (%), and false alarm rate (FAR %) for severe weather warnings for tornadoes.** For FY 2001, NWS met one of the three performance targets for this measure—false alarm rate (73 percent). But more significant is the fact that as currently calculated, the measure does not convey that for 43 percent of the tornadoes recorded nationally during the year, NWS issued either a warning with no lead time or no warning at all. The lead time of 10 minutes reported in the Department's *FY 2001 Annual Program Performance Report/FY 2003 Annual Performance Plan* and the *FY 2001 Accountability Report* represents a national average.³ In addition, we found that internal controls over the measure do not preclude the reporting of incorrect data, and that procedures for maintaining documentation are inconsistent. Also, problematic is the presentation of the measures in the *FY 2001 APPR/FY 2003 APP* and the *FY 2001 Accountability Report*: key details—such as not explaining the use of estimates in some cases and the exclusion of data from certain areas in others—limit the usefulness of the results. Finally, we found that certain data quality control procedures were not initially applied to all FY 2002 data—a condition that would have limited the reliability of the results. However, in response to our concern about this issue, NWS implemented an interim quality check, pending upgrades to the Advanced Weather Interactive Processing System (AWIPS), which should include a permanent data-checking feature.

NOAA and NWS need to (1) augment the current measure to reflect the percentage of tornadoes for which the public is warned with no lead time or not warned at all; (2) strengthen internal controls, including requirements for supporting documentation; (3) implement procedures that ensure the reporting of the most accurate data; and (4) enhance the discussion of performance results. (See page 7.)

² *National Oceanic and Atmospheric Administration: Financial Statements, Fiscal Year 2001*. Audit Report No. FSD-14475-2-0002/February 2002.

³ For FY 2001, the public was provided with no lead time 43% of the time, less than 10 minutes of lead time, 18% of the time, and equal to or greater than 10 minutes of lead time 39% of the time.

2. **Measure: Lead time (minutes) and accuracy (%) for severe weather warnings for flash floods.** While NWS met both performance targets for this measure, we found that—as with tornadoes—FY 2001 results do not convey that for 27.2 percent of the flash floods recorded nationally, NWS issued either a warning with no lead time or no warning at all. The lead time of 46 minutes reported in the Department's *FY 2001 APPR/FY 2003 APP* and the *FY 2001 Accountability Report* represents a national average.⁴ Also, we found instances in which flash flood data incorrectly included data related to other types of flooding and lead times and event times were inaccurately recorded. And again, we found that NOAA's presentation of the measures in the *FY 2001 APPR/FY 2003 APP* and the *FY 2001 Accountability Report* needed additional disclosures such as the use of estimates in some cases and the exclusion of data from certain areas in others. Similar to the measure related to tornadoes, the initial decision not to employ a type of quality control procedure threatened the reliability of FY 2002 data. For this measure, we made the same recommendations as for the tornado measure: that NOAA and NWS (1) augment the current measure to reflect the percentage of flash floods for which the public is warned with no lead time or not warned at all; (2) strengthen internal controls, including requirements for supporting documentation; (3) implement procedures that ensure the reporting of the most accurate data; and (4) enhance the discussion of performance results. (See page 12.)
3. **Measure: Accuracy (%) of three-day forecast of precipitation.** This measure—designed to gauge how well NWS accurately predicts precipitation 3 days in advance—does not fully reflect performance because it does not take into account areas where precipitation is forecasted but does not occur. This data limitation is not disclosed in the *FY 2001 APPR/FY 2003 APP*; neither are data verification procedures nor the fact that the forecasts counted are for 1 inch or more of rain. NOAA should (1) note in future performance reports that FY 2001 results did not measure areas where precipitation was forecasted but did not occur; (2) state that the measure only counts forecasts of 1 inch or more of precipitation; and (3) give greater detail about verification procedures. (See page 17.)

Performance Goal: Implement seasonal to interannual climate forecasts

1. **Measure: Determine the accuracy of the correlation between forecasts of the Southern Oscillation Index (SOI) and El Niño/La Niña events.** This measure purports to assess the accuracy of the correlation between forecasts of South Pacific sea surface temperature and actual sea surface temperature of the waters connecting Darwin, Australia, and Tahiti. However, the measure does not demonstrate accuracy of the correlation, but rather proximity—that is, how closely NWS forecasts correlate with

⁴ For FY 2001, the public was provided with no lead time 27% of the time, less than 46 minutes of lead time, 37% of the time, and equal to or greater than 46 minutes of lead time, 36% of the time.

observed temperatures. Therefore, the measure—as titled and as described in the *APPR/APP*—does not clearly articulate what is being assessed, and does not explain the correlation index or that the measure is based on cumulative data. We recommend that the performance measure be either eliminated or revised. If the latter option is chosen, NOAA should provide the necessary disclosures and explanations of changes in subsequent performance reports. (See page 20.)

2. **Measure: U.S. temperature forecasts (skill score).** This measure reflects NWS's success at accurately predicting temperature over the prior 48 months. We found that while performance results for FY 2001 were calculated using a 48-month seasonal average, results for fiscal years 1999 and 2000 were based on 36 months, which rendered FY 2001 data noncomparable with that from these earlier years. And, as with other measures, NOAA excludes key details from the measure's discussion in the *APPR/APP* and Accountability Report, in this case that the measure is a cumulative average, covers forecasts for less than half of the U.S, and is calculated manually. Neither does NOAA explain the reported score's decline from FY 2000 to FY 2001—detail that would likely be of interest to decision makers and the public. NOAA should revise reported results for FYs 1999 and 2000 using the 48-month average and explain the change in subsequent APPRs; disclose that the average is cumulative; and develop and enforce procedures for ensuring that data is verified before it is published. (See page 22.)
3. **Measure: Number of new monitoring or forecast products that become operational per year (cumulative).** This measure reports on the development of new products for monitoring weather or issuing forecasts. We found that NWS has no clear definition of what constitutes a new product or formal procedures for verifying the numbers of new products reported by the office that supplies this information—the National Climatic Data Center. NOAA reported in the *FY 2001 APPR/FY 2003 APP* that it introduced 4 new products—its target number. However, our audit found that, depending on the definition used, 50 new products could have been identified as becoming operational in FY 2001. NWS needs to develop such definitions and procedures to avoid confusion over what is a new product. (See page 25.)
4. **Measure: New climate observations introduced.** This measure is intended to record the number of new monitoring systems deployed or made operational during the fiscal year. The title of the new measure and its accompanying narrative in the *FY 2001 APPR/FY 2003 APP* imply that NOAA is counting multiple climate observation systems or products, when in reality it is measuring the introduction of only Argo floats—free-drifting floats that gauge temperature and salinity of the upper 2,000 meters of the ocean. NOAA's FY 2001 target for this measure was 120 floats, and it reported introducing 132. However, this number actually identifies the number of floats *budgeted for procurement* in the fiscal year, not deployed.

If NOAA plans to report only on Argo float deployments, it should revise the measure and discussion accordingly and count actual deployments. If not, NOAA should include all new observation equipment deployed. (See page 27.)

The accuracy and reliability of reported performance measures is largely a function of adequate internal controls. NOAA management is responsible for implementing such mechanisms and ensuring the quality of reported information. Therefore, we believe that NOAA should promptly correct the identified internal control weaknesses that led to inaccurate, incomplete, or unclear performance reporting, and thereby enhance the credibility and usefulness of performance results for Congress, OMB, and other stakeholders.

In responding to the draft report, NOAA either concurred with or is taking corrective action consistent with all of the recommendations. Within its response, NOAA identified corrective actions taken or planned and implementation schedules. However, NOAA expressed concern that (1) the tone and tenor of the report was overly negative relative to the findings and recommendations presented and (2) the sample size used to examine performance data for tornado and flash flood lead times was not large enough to characterize identified deficiencies as internal control weaknesses. Nevertheless, the NWS response includes planned actions to strengthen internal controls and reduce the likelihood of the kinds of lapses in internal controls identified during our review. NOAA also suggested that we include additional discussion of the process used by the agency to ensure the accuracy of performance information in the report.

We are encouraged by the actions taken or planned by NOAA. Where appropriate, we have modified the report to reflect NOAA's response. However, we believe the tone and tenor of the final report is consistent with the need for NOAA to improve the completeness, accuracy, and clarity of performance information it reports. Within the appropriate sections of this report, we summarize NOAA's response to our draft report as well as provide our comments. NOAA's complete response is attached to the report as Appendix I.

This is the second report issued on NOAA's performance measures. The first—*Improvements Needed in the Reporting of Performance Measures Related to Promoting Safe Navigation and Sustaining Healthy Coasts* (FSD-14998-3-0001)—was issued in February 2003.

INTRODUCTION

The National Oceanic and Atmospheric Administration (NOAA) is charged with assessing and predicting changes in the Earth's environment, and protecting and managing marine and coastal resources to ensure sustainable economic opportunities—missions that support one of the Department of Commerce's three strategic goals: "Observe and manage the Earth's environment to promote sustainable growth."⁵

Both NOAA and the Department report on the performance of NOAA programs and activities to meet the requirements of the Government Performance and Results Act of 1993 (GPRA). GPRA seeks to improve the effectiveness, efficiency, and accountability of federal programs by requiring agencies to set performance goals and to annually compare actual performance against those goals and report the results. The Department presented performance information for goals and measures considered to be critical in its *FY 2001 Accountability Report*, and in the more recent *FY 2002 Performance and Accountability Report*.

NOAA has seven goals against which to assess and report on its program and financial performance:

- Build sustainable fisheries.
- Sustain healthy coasts.
- Recover protected species.
- Advance short-term warnings and forecasts.
- Implement seasonal to interannual climate forecasts.
- Predict and assess decadal to centennial climate change.
- Promote safe navigation.

Within each goal are measures that NOAA uses to assess the programs and activities of its five line offices: the National Ocean Service; National Marine Fisheries Service; Office of Oceanic and Atmospheric Research; National Weather Service (NWS); and National Environmental Satellite, Data, and Information Service (NESDIS). From this assessment, NOAA generates performance results to enable Congress, the Office of Management and Budget (OMB), and other decision makers to evaluate the federal government's investment in these programs, and help agency officials improve program outcomes. However, performance results support these objectives only to the extent that the data is reliable. GPRA requires agencies to verify and validate performance data to provide assurance of its reliability. The General Accounting Office defines verification as the "assessment of data completeness, accuracy, and consistency, and the related quality control practices," and validation as the "assessment of whether the data is appropriate for the performance measure."⁶

⁵ The Department's other two strategic goals are (1) to provide the information and framework to enable the economy to operate efficiently and equitably, and (2) to provide infrastructure for innovation to enhance American competitiveness.

⁶ U.S. General Accounting Office, July 30, 1999. *Performance Plans: Selected Approaches for Verification and Validation of Agency Performance Information*, GAO/GGD-99-139. Washington, DC: U.S. General Accounting Office.

Tying Costs to Results

NOAA's audited FY 2001 financial statements break down the bureau's total \$2.71 billion net costs of operations by performance goal, as follows: build sustainable fisheries--\$493 million; sustain healthy coasts--\$285 million; recover protected species--\$164 million; advance short-term warning and forecast services--\$1.43 billion; implement seasonal to interannual climate forecasts--\$116 million; predict and assess decadal to centennial climate forecasts--\$102 million; and promote safe navigation--\$120 million.⁷ The Department provided FY 2001 enacted budget amounts by goal in its *FY 2001 Annual Program Performance Report (APPR)* and *FY 2003 Annual Performance Plan (APP)*, which also details departmental and bureau efforts to comply with GPRA. Together, the two sources permit analysis of FY 2001 performance results—by goal—in terms of the federal government's financial investment in achieving them. Depending on the extent to which decision makers rely on performance information in allocating resources, the credibility of reported data may affect the amounts ultimately budgeted for a specific program.

This report details our audit of two NOAA goals and a selection of their associated measures:

Advance short-term warnings and forecasts

1. Lead time (minutes), accuracy (%), and false alarm rate (FAR, %) for severe weather warnings for tornadoes
2. Lead time (minutes) and accuracy (%) for severe weather warnings for flash floods
3. Accuracy (%) of three-day forecast of precipitation

Implement seasonal to interannual climate forecasts

1. Determine the accuracy of the correlation between forecasts of the southern oscillation index (SOI) and El Niño/La Niña Events
2. U.S. temperature forecasts (skill score)
3. Number of new monitoring or forecast products that become operational per year (cumulative)
4. New climate observations introduced

These two goals are primarily supported by the programs and activities of the National Weather Service; National Environmental Satellite, Data, and Information Service; and the Office of Oceanic and Atmospheric Research. Over the past several years, NWS has received accolades for its collection and reporting of performance results. For example,

⁷ *National Oceanic and Atmospheric Administration: Financial Statements, Fiscal Year 2001*. Audit Report No. FSD-14475-2-0002, February 2002.

the April 2001 issue of *Government Executive*⁸ magazine described NWS as having “[r]esults focused management coordinated throughout the agency and across functions to achieve mission success.” GAO—in its *Observations on the Department of Commerce’s Fiscal Year 1999 Annual Program Performance Report and Fiscal Year 2001 Annual Performance Plan* (June 30, 2000)—stated that “NOAA’s existing and new measures were quantifiable and outcome-oriented and provided a succinct statement of expected performance.” In an August 1999 letter to the Secretary of Commerce, the chairman of the Senate Governmental Affairs Committee noted that “NWS established goals and targeted performance levels that balance the competing priorities of increasing both the lead times and accuracy of severe warnings.” While these reviews appropriately recognized NOAA’s commitment to reporting outcome-oriented measures, our review involved a comprehensive review of NOAA internal controls in place to ensure the reporting of accurate and reliable performance data.

The NWS informed us that it dedicates significant effort to ensuring the accuracy of performance information. For example, the NWS conducts site visits of regional offices and weather forecast offices and maintains documentation to verify its ability to warn the public. Also, a centralized tracking data base for performance information with automated quality controls is maintained.

NWS’ Performance Branch (within its Performance and Awareness Division) is responsible for verifying the accuracy and timeliness of warnings, forecasts, and all other NWS services, and for promoting the importance of verification throughout the line office. NOAA uses the verified data as a baseline for establishing GPRA-mandated performance measures.

This audit report, the second on NOAA goals and measures,⁹ details our findings and recommendations regarding NOAA procedures for collecting, verifying, and presenting performance data.

⁸ In its April 2001 issue, *Government Executive Magazine* gave NWS straight As in the areas of financial, human resource, information, and physical assets management, noting its “results-focused management coordinated throughout agency and across functions and to achieve mission success.”

⁹ *National Oceanic and Atmospheric Administration: Improvements Needed in the Reporting of Performance Measures Related to Promoting Safe Navigation and Sustaining Healthy Coasts*, Audit Report No. FSD-14998-3-0001, February 2003.

OBJECTIVES, SCOPE, AND METHODOLOGY

Our purpose was to (1) assess the collection and reporting of NOAA performance information in documents submitted to meet GPRA requirements, and (2) determine whether NOAA's internal controls are sufficient to ensure that performance data is accurate, consistent, and reliable.

To evaluate the measures, the reliability of reported results, and the usefulness of performance information, we did the following:

- Reviewed federal guidance and legislation, including GPRA; the CFO Act; OMB Circular A-123, *Management Accountability and Control*; OMB Circular A-11, *Preparation, Submission, and Execution of the Budget*, Part 2; and GAO *Standards for Internal Control in the Federal Government*.
- Toured the weather forecast office in Sterling, Virginia.
- Interviewed NOAA officials responsible for generating, maintaining, and reporting performance data.
- Identified and tested internal controls.
- Subjected the data to validation and verification procedures.
- Evaluated the clarity and usefulness of explanations provided for each measure in the FY 2001 *Accountability Report* and the FY 2001 APPR/FY 2003 APP.
- Performed a cursory review of the FY 2002 *Performance and Accountability Report* to see what, if any, actions NOAA had already taken to address concerns we raised during the course of our review.

We further tailored our audit procedures to each measure under review, as follows:

- **Lead time (minutes), accuracy (%), and false alarm rate (FAR, %) for severe weather warnings for tornadoes.** We selected a judgmental sample of 20 weather forecast offices, and for each, assessed all the warnings issued and tornadoes verified during FY 2001 and the first 6 months of FY 2002 and evaluated selected performance statistics for lead time, accuracy, and false alarm rate. Also, we reviewed supporting documentation for 20 tornadoes. From this sample, we noticed that the public frequently received little or no lead time with tornado warnings. Consequently, we obtained a national listing of all FY-2001 tornado events for which the public received warnings with lead times to determine the frequency in which the public was provided little or no time to take precautionary actions.
- **Lead time (minutes) and accuracy (%) for severe weather warnings for flash floods.** Again, using a judgmental sample of 20 weather forecast offices, we assessed for each office, all the warnings that were issued and verified during FY 2001 and the first 6 months of FY 2002; evaluated selected performance statistics; and reviewed supporting documentation for 20 flash floods. Again, from our sample, we noticed that the public frequently received little or no lead time with flash flood warnings. Consequently, we also obtained a national listing of all FY 2001 events for which the public received warnings with lead times to

determine the frequency in which the public was provided little or no time to take precautionary actions.

- **Accuracy (%) of three-day forecast of precipitation.** We selected a judgmental sample of daily precipitation data for the months of June, July, August, and September 2001. We compared the results of observed precipitation for each month and the correct forecasts with the overall monthly totals used to calculate the measure.
- **Determine the accuracy of the correlation between forecasts of the southern oscillation index (SOI) and El Niño/La Niña events.** We independently verified the correlation between the forecast and actual observations of sea surface temperature for the past 16 years beginning October 1985 and ending September 2001, and compared our findings with the reported results of FY 2001.
- **U.S. temperature forecasts (skill score).** We selected a sample of quarterly reports containing skill scores (that is, forecast accuracy rates) for the 5 seasons that spanned July 2000 to September 2001, and compared the scores in our sample with those used to calculate the cumulative 48-month average score reported in the *FY 2001 APPR/FY 2003APP*. We also evaluated and verified the FY 1999 and FY 2000 skill scores reported in the *APPR/APP*.
- **Number of new monitoring or forecast products that become operational per year (cumulative).** We interviewed meteorologists from NWS' Climate Prediction Center (CPC) who develop new monitoring or forecast products, and compared the number of new products they identified as becoming operational in FY 2001 with the results reported in the *APPR/APP*.
- **New climate observations introduced.** We asked NOAA to tally the number of Argo floats (marine temperature/salinity gauges) successfully deployed in FY 2001, as well as the type and number of other weather observation equipment put in place, and compared this number against FY 2001 reported results, which pertain to Argo floats only.

We did not test the reliability of computer-generated data for the performance measures as such data was not essential to satisfying our audit objectives. For the measures relating to tornadoes and flash floods, we obtained – through interviews – a high level of understanding of the data integrity controls over the information systems used to collect and report this information, and noted nothing of concern regarding the credibility of that data. Neither was it our purpose to determine whether these performance measures are the most appropriate for the bureau.

We conducted our fieldwork from September 2002 to March 2003 at Camp Springs, Maryland, and NOAA headquarters in Silver Spring, Maryland, and performed this audit in accordance with *Government Auditing Standards* issued by the Comptroller General of the United States, and under authority of the Inspector General Act of 1978, as amended, and Department Organization Order 10-13, dated May 22, 1980, as amended.

FINDINGS AND RECOMMENDATIONS

In general, we found that the seven performance measures we reviewed need stronger internal controls to better ensure the accuracy and reliability of reported results, as well as some revision and additional disclosures to make the results more meaningful and useful. While we found a commitment on the part of NOAA to report outcome-oriented measures and reliable information, we found that (1) reported performance data at times did not provide a complete picture of performance; (2) performance information was not always accurately reported, and (3) explanatory language frequently did not appropriately describe results or the limitations of the data.

These issues are consistent with those identified during our reviews of performance results within NOAA and the rest of the Department. For the measures that assess tornado and flash flood warnings, we identified similar overall weaknesses and thus offered the same recommendations.

Guidance on maintaining internal controls and reporting performance information is contained in the following:

- The Government Performance and Results Act of 1993, Section 4, states that each agency establish performance indicators to be used in measuring the relevant outputs, service levels, and outcomes of each program activity.
- OMB Circular A-11, Part 2, *Preparation, Submission, and Execution of the Budget*, requires agencies to include in their annual plans a description of how they intend to verify and validate actual performance. The methods "should be sufficiently credible and specific to support the general accuracy and reliability of the performance information that is recorded, collected, and reported."¹⁰
- OMB Circular A-123, *Management Accountability and Control*, identifies internal controls as the organization, policies, and procedures used by agencies to reasonably ensure that reliable and timely information is obtained, maintained, reported, and used for decision making. Section II states, "documentation for transactions, management controls, and other significant events must be clear and readily available for examination."
- *GAO Standards for Internal Control in the Federal Government* states that an agency's control activities must ensure that all transactions are completely and accurately recorded.

¹⁰ In June 2002, OMB Circular A-11 was revised. Identical language is now contained in Part 6, *Preparation and Submission of Strategic Plans, Annual Performance Plans, and Annual Program Performance Reports*, Section 220.5.

I. Performance Measure: Lead Time (Minutes), Accuracy (%), and False Alarm Rate¹¹ (FAR, %) for Severe Weather Warnings for Tornadoes

Measurements of NWS' track record for issuing accurate, timely warnings as well as false warnings provides a robust assessment of its performance in warning the public about impending tornadoes. Moreover, NWS informed us that it makes significant efforts to ensure the accuracy and reliability of the performance information it reports. However, we found that NWS needs to expand the current measure to provide a more complete picture of performance, strengthen its internal controls, including requirements for supporting documentation, provide additional explanations of the data, and take steps to ensure that the data is as accurate as possible.

For FY 2001, NWS met the performance target for false alarm rate, which was 73 percent. It missed its target for average warning lead time by 3 minutes (reporting a 10 minute average as opposed to the goal of 13 minutes), and its target for accuracy by one percentage point (achieving 67 percent as opposed to 68 percent)—a difference the agency maintains is statistically insignificant and well within the standard deviation for the measure.

NWS assesses its ability to warn the public about tornadoes by measuring lead time (minutes), accuracy (%), and false alarm rates (FAR %) of issued warnings.

- Lead time is the time that elapses between when a warning is issued and the tornado strikes. The reported performance statistic is the average of lead times for all tornadoes that occurred during the fiscal year.
- Accuracy is the percentage of times a tornado actually occurs in the area covered by the warning.
- False alarm rate is the percentage of times a tornado warning is issued but not verified as having occurred.



Picture source: NOAA

We evaluated a judgmentally selected sample of 20 events—including warnings issued, supporting documentation, and related outcomes and events—recorded by 20 of the 116 weather forecast offices (WFOs) that report on this measure.¹² (See Table 1.)

A. Average lead time does not convey the full picture of NWS performance.

This measure is a useful indicator of performance, but does not fully reflect NWS' track record for providing warnings about impending tornadoes because it does not convey the many instances in which the public receives no advance warning—that is, it has no time to act—or receives no warning at all.

¹¹ The false alarm rate was added as a reportable measure in FY 2000, although this data had been collected and used internally previously.

¹² There are a total of 123 WFOs, but only 116 feed this measure.

Average annual lead times for the 20 WFOs in our sample ranged from 24.5 minutes to 0 minutes, a significant disparity. For the 321 tornadoes that occurred in areas covered by these offices, the public received no lead time to take precautionary action in 91 instances, or 28 percent of the time.

Nationally, we found that for the 1,205 tornadoes recorded in FY 2001

- the public received warning with lead times on only 691 occasions, or 57 percent of the time;
- for 32.5 percent of these occasions, the public received no warning, and for the remaining 10.5 percent of the recorded events, the public was warned without any lead time;
- the public was warned with lead times between 1 and less than 10 minutes 18 percent of the time; and
- the public received warnings equal to or more than 10 minutes of lead time 39 percent of the time.

One very long lead time will have a greater impact on the average than will several lead times of zero. For example, in a number of instances nationally, lead times were longer than 40 minutes. Hence, averaging the performance data for a function that has such across-the-board variation may not completely portray how effective NWS is at providing warnings because the average masks the large number of events for which

there are inadequate lead times, no lead times, or no warnings at all.

TABLE 1. Weather Forecast Offices Evaluated for Tornado Warning Response	
Austin, TX	Morehead City, NC
Billings, MT	Northern, IN
Bismarck, ND	Philadelphia, PA
State College, PA	Pueblo, CO
Eastern, ND	Raleigh, NC
Glasgow, MT	Salt Lake City, UT
Jackson, KY	Tampa Bay Area, FL
Las Vegas, NV	Tulsa, OK
Little Rock, AR	St. Louis, MO
Midland, TX	Wakefield, VA

One approach the NWS could take to enhance the usefulness of the results for this measure would be to provide percentages of instances in which it failed to issue warnings with lead times or to issue any warning at all, and to set a goal to reduce those percentages. These statistics can be derived from data the agency currently collects.

B. Internal controls over generating and reporting data should be strengthened.

The FY 2002 performance results for false alarm rate and accuracy contained in the *FY 2002 Performance and Accountability Report* were incorrectly reported: the false alarm rate was given as 76 percent, when the correct rate was 73 percent; and the accuracy rate was given as 77 percent instead of 76 percent. The misstatement occurred as information was updated and transferred from NOAA headquarters to the Department, and the mistake was not identified until after the report was issued.

Except for one instance, the supporting documentation for the one event per office we reviewed was generally complete and consistent with information reported, as well as with tornado data maintained in NWS' database: in the one exception, the event time for a tornado reported by a WFO was incorrectly entered into the database, which resulted in a 10-minute lead time being inaccurately assigned to this event. The error increased the

average lead time for all tornado warnings at this WFO for FY 2001 to 2.5 minutes when it should have been 0 minutes. While this deficiency did not impact national results for FY 2001, it revealed a potential internal control weakness, thus leaving open the possibility for material errors (e.g., significant variances) at specific offices. If the entry of incorrect data into the database became commonplace as opposed to the exception, the credibility of national results could be adversely impacted.

Consequently, NWS needs to implement procedures to guard against the potential for material error, such as (1) conducting independent spot checks of supporting documentation against reported data at the WFO, regional offices, or headquarters, or (2) requiring attestations from each WFO meteorologist-in-charge as to the accuracy of data submitted by the office.

C. Consistent procedures for maintaining supporting documentation need to be established and enforced.

Two WFOs could not readily provide documentation we requested, and we found that the offices were uncertain about how long they should maintain such records, largely because of conflicting guidance. One office explained that it retains actual tornado warnings on file for "a limited period." Another office stated that—consistent with the *Weather Service Operations Manual*, Chapter C-40, section 5.57—it retains records for 2 years. WFO staff told us that NWS has been revising its policies and procedures, but that new guidelines do not specify how long the offices should retain documentation. Thus, the inconsistency in available documentation persists.

D. Additional disclosures would enhance the usefulness of reported results.

NOAA's presentation of this measure implies that it represents data from the entire country, when in fact it does not include results for Hawaii and Puerto Rico. Hawaii and Puerto Rico are not included because in the past NWS did not have the systems capability to capture this data and the infrequency of events in these locations. We believe, however, that decisionmakers/stakeholders should be made aware that reported results are not all-inclusive. In addition, the NWS frequently uses estimates as opposed to the actual times a tornado occurs, a fact that is not disclosed.

NWS states that time estimates are necessary and reasonable when spotters are unavailable or times reported by various sources conflict. In these situations, NWS arrives at an estimate based on past experience and available weather forecast data. We agree that the use of estimates is reasonable under certain circumstances, but that their use should be clearly disclosed so that the reader fully understands the limitations associated with reported results.

E. Not implementing certain quality control procedures could have impacted reliability.

NWS' Performance Branch at one time had a staff member who provided quality control checks of performance data submitted by the WFOs. However, the quantity of data

became too voluminous for a single individual to handle, and in December 2001, NWS suspended the checks indefinitely.

In transitioning from manual to automated quality controls, it developed two automated procedures, referred to as Rule 1 and Rule 2, to ensure accuracy of performance information and provide WFOs with real-time feedback when errors are detected in warnings. Rule 1, an automated procedure to check for warning coding and format errors, was implemented in January 2002. Rule 2, the second of two temporary data modifications pending upgrades to the Advanced Weather Interactive Processing System (AWIPS) to screen out inappropriate data and thus improve quality control, was not implemented. After much debate about the potential effects of Rule 2 on performance results, NWS decided not to implement it and thus provided unedited preliminary numbers to NOAA's chief financial officer. We brought this to the attention of the deputy chief financial officer for NWS. NWS subsequently implemented Rule 2.

We commend NWS management for promptly addressing the need for additional quality control procedures. NWS informed us that plans are in place to upgrade AWIPS software to screen out inappropriate data. Specifically, new software to be released in December 2003 will allow the tracking of warnings by event number and identify warnings as new or corrected. Also, quality control checks have been added to the software that generates warnings. These checks will include safeguards to ensure the proper coding and formatting of warnings. We believe that, consistent with other priorities, the software should be upgraded to ensure the data collected is accurate as well.

F. Recommendations

The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA does the following:

- (1) Provides performance data to reflect the percentage of events in which the public is not provided a warning in time to take precautionary actions.
- (2) Strengthens internal controls over performance reporting, to include a policy for maintaining accurate supporting documentation and procedures for reconciling data prior to reporting.
- (3) Revises the presentation of the measure to include all appropriate disclosures to make the discussion of results more clear and meaningful.
- (4) Takes other appropriate actions—such as upgrading systems and software—to facilitate and ensure the reporting of accurate data.

G. NOAA Response

In response to the draft report, NOAA concurred with all four recommendations, identified corrective actions taken or planned, and provided target completions dates.

NWS will develop a new performance measure that augments the current set of performance measures. In addition, both NOAA Headquarters and NWS will strengthen internal controls to reduce the likelihood of reporting inaccurate performance data. NWS will improve the disclosures of quality and sources of data in future reports containing performance information. Also, the NWS provided information on future efforts to automate quality controls of its performance data. A software upgrade, scheduled for December 2003, will include a capability to allow the improved tracking of warnings. In addition, quality control checks to ensure the proper coding and formatting of warnings have been added to the software that generates warnings.

In the response, NWS expressed concern over the sample size used to examine performance data for tornado lead times and recommended we modify the discussion of agency efforts to implement quality controls over the performance data. NWS described the sample size to identify an internal control weaknesses as small and not representative. Nevertheless, NWS stated that it will reemphasize the need for accurate data and develop a process to spot-check data for accuracy. NWS recommended that we reference the implementation of an automated procedure to check for warning coding and format errors, Rule 1, in January 2002 and qualify the wording regarding the suspension of quality control checks.

H. OIG Comments

We commend NOAA and the NWS for the corrective actions taken and planned. We believe these actions demonstrate a commitment to the reporting of reliable performance data. Where appropriate, we have modified the report to reflect NOAA's response. The modifications included: (1) discussing the NWS implementation of Rule 1 in January 2002, (2) clarifying the initial decision not to implement Rule 2, and (3) providing additional discussion of planned automated quality control upgrades. With respect to the NWS concern over sample size, we believe the report makes it clear that the identified problem was an exception and that we do not project this error over the population of tornado events.

II. Performance Measure: Lead Time (Minutes) and Accuracy (%) for Severe Weather Warnings for Flash Floods

For FY 2001, NWS met its accuracy goal of 86 percent, and exceeded its lead time goal of 45 minutes by 1 minute. NOAA indicated in the *FY 2001 APPR/FY 2003 APP* that performance results tend to be higher when the number of events is above average in a given year.

We reviewed a judgmentally selected batch of warnings, subsequent events, and related supporting documentation from 20 of the 116 WFOs that report for this measure (Table 2).

A. Average lead time does not fully capture performance.

As with tornadoes, we believe that averaging lead times to quantify performance does not present a fully accurate picture of NWS' track record for warning the public about impending flash floods because it does not convey the many instances in which people are either warned without lead time or not warned at all.

- Average annual lead times for the 20 WFOs in our sample ranged from 0 to 228 minutes, a significant disparity. For the 780 flash flood events identified as having occurred in areas covered by these offices, the public received no lead time in 164 instances, or 21 percent of the time.

Nationally, we found that for the 2,779 flash floods recorded in FY2001, the public received either no warning or a warning with no lead time 27 percent of the time: the public received between 1 and less than 46 minutes of lead time, 37 percent of the time, and equal to or greater than 46 minutes of lead time, 36 percent of the time.

NWS assesses its ability to warn the public about flash floods by measuring lead time (minutes) and accuracy (%) of issued warnings.

- Lead time is the time that elapses between when the warning is issued and the flash flood strikes. To get this statistic, NWS averages the lead times for all flash floods that occur during a given year.

- Accuracy is the percentage of times a flash flood actually occurred in the warning area.



Flash flood races across road

Picture source: NOAA

TABLE 2: Weather Forecast Offices Evaluated for Flash Flood Warning Response

Albuquerque, NM	Flagstaff, AZ
Amarillo, TX	Grand Rapids, MI
Billings, MT	Great Falls, MT
Birmingham, AL	Kansas City, MO
Boston, MA	Morehead City, NC
State College, PA	Nashville, TN
Wilmington, OH	Philadelphia, PA
Quad Cities, IA	Portland, ME
Dodge City, KS	Riverton, WY
Eastern, ND	Wakefield, VA

On a number of occasions, the WFOs provided warnings with lead times of more than 5 hours, which—when factored in with other, shorter lead times—skews the average upward, thus masking the many instances in which there were no warnings, no lead times, or lead times of only a few minutes.

Consequently, NWS could enhance its presentation of results by providing the percentages of instances in which flash floods occurred without any lead time or with no warning, and should set as a goal, its success at reducing these percentages. The required statistics can be derived from data already being collected by NWS.

B. Internal controls should be strengthened.

As is the case with tornado warning data, NWS headquarters has personnel and procedures for verifying flash flood information. However, we identified data gathering weaknesses that suggest the need to improve internal controls. We note that the WFOs should be the first line in the verification process—ensuring accuracy by diligently determining whether and when an event occurred and properly documenting their observations, as the data they report impacts all statistics provided under the measure. Beyond that, NWS must augment current internal controls to eliminate the following problems:

- **Mixing of flood data in the flash flood database.** NWS defines a flash flood as a flood that occurs within 6 hours of the causal event (e.g., rainfall). The performance measure is meant to include only flash floods; however, we identified 4 instances of 20 in which WFOs counted regular floods as flash floods. Three of the four events occurred in NWS' Eastern Region, where we understand categorizing between floods and flash floods has been an issue for many years. The mixing of flood and flash flood data—and the inconsistencies it suggests in data gathering among the regions—undermines the reliability of nationally reported performance results for this weather event. NWS informed us that on August 6, 2001, it issued a policy prohibiting issuance of flash flood warnings for general area flooding. NWS further stated that with the issuance of NWS Instruction 10-1605 on January 6, 2003, NWS established a policy to include only flash flood warnings and event in the NWS flash flood verification program.
- **Incorrect event times.** After NWS issues a warning, it tries to determine whether the event in fact occurred and, if so, how much advance warning time the public received. We found that one WFO recorded an event as occurring at 4:30 p.m., while supporting documentation (confirmed in discussions with WFO staff) indicated that the event actually occurred at 4:25 p.m.—a loss of 5 minutes in lead time.
- **Incorrect lead times.** We identified one instance in which the WFO tied a single flood event to two prior flash flood warnings, and thus incorrectly reported lead times of 4 minutes and 179 minutes, when in reality it was zero.

- **Warning mistakenly issued.** A WFO recycled a previously issued warning but failed to change the warning area designated on the original warning document to the area covered by the subsequent event. As a result, counties not at risk for the severe weather event were erroneously notified of its approach, and the WFO's false alarm rate was negatively impacted.

While none of these internal control deficiencies significantly impacted nationally reported results, they do leave open the possibility for material errors in the future. NWS must implement procedures to guard against this possibility.

C. Consistent procedures for maintaining supporting documentation need to be established and enforced.

As with tornado warnings, we found that supporting documentation for flash floods was generally complete and consistent with reported information, but that there was uncertainty among the WFOs about how long to maintain supporting documentation (i.e., 1 year or 2), and insufficient procedures for storing and later retrieving electronic records. As a result, the records we requested for two events were not readily available. The confusion appears to be prompted by the anticipated revision of NWS procedures. Current guidance (see *Weather Service Operational Manual*, Chapter C-40, section 5.57) calls for records to be retained for 2 years. NWS personnel informed us that the new guidelines, still in draft form, do not specify the length of time offices should retain documentation.

D. Additional disclosures would enhance the usefulness of reported results.

As was the case with tornado warnings, the presentation of this measure implies that it represents data from the entire U.S., when in fact it does not include results for Alaska, Hawaii, Puerto Rico, Guam, and U.S. Virgin Islands, and frequently uses estimates of event times.

NWS reportedly excludes these states and territories because it does not have the communications capability to gather this data from them, and because of the historical infrequency of flash floods in these locations. Alaska, Hawaii, and Puerto Rico, however, had a significant number of flash floods during FY 2001. Regardless, the reader should be made aware that reported results are not all-inclusive, and NWS should specify the excluded areas.

NOAA states that time estimates are necessary and reasonable when spotters are unavailable or times reported by various sources conflict. In these situations, NOAA develops an estimate based on past experience and available weather forecast data. We agree that the use of estimates is reasonable under certain circumstances. However, we believe NOAA should clearly state when estimates are used in place of actual times, so that the reader fully understands the limitations associated with the reported results.

E. Not implementing certain quality control procedures could have impacted reliability.

NWS' decision not to implement certain quality control checks in December 2001 (see page 10), impacted flash flood data collection as well. NWS implemented Rule 1, an automated procedure to check for warning coding and format errors, in January 2002. With the implementation of Rule 2, before the final FY 2002 numbers were calculated, the FY 2002 results reported for flash flood lead times were corrected with a reduction of 4 minutes. We are encouraged by the prompt action NOAA management took to restore quality control procedures.

As noted earlier, plans are in place to upgrade AWIPS software to screen out inappropriate data. We believe that, consistent with other priorities, the software should be upgraded to ensure accurate data as well.

F. Recommendations

The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA does the following:

- (1) Provide performance data to reflect the percentage of events in which the public is not provided with a warning in time to take precautionary actions.
- (2) Strengthen internal controls for the performance measure, to include a policy for maintaining accurate supporting documentation and procedures for reconciling data prior to reporting.
- (3) Revise the presentation and all appropriate disclosures to make the discussion of results more clear and more meaningful.
- (4) Take other appropriate action—such as upgrading systems and software—to facilitate and ensure the reporting of accurate data.

G. NOAA Response

In response to the draft report, NOAA concurred with all four recommendations, identifying corrective actions taken or planned and target completions dates. NWS stated that it will develop a new performance measure that augments the current set of performance measures. NWS will strengthen internal controls to reduce the likelihood of reporting inaccurate performance data. Also, NWS will improve the disclosures of quality and sources of data in future reports containing performance information. Additionally, NWS provided information on future efforts to automate quality controls of its performance data. The NWS noted that a software upgrade, scheduled for December 2003, will include a capability to allow the improved tracking of warnings. In addition,

quality control checks to ensure the proper coding and formatting of warnings have been added to the software that generates warnings.

Also, in the response, NWS stated that it had already taken actions to strengthen its internal controls to ensure the reporting of accurate data. NWS stated that it had already implemented a policy prohibiting issuance of flash flood warnings for general area flooding on August 6, 2001. On January 6, 2003, NWS issued a policy to include only flash flood warnings and events in the NWS flash flood verification program. Also, NWS recommended that we reference the implementation of Rule 1, an automated procedure to check for warning coding and format errors in January 2002 and qualify the wording regarding the suspension of quality control checks by NWS.

H. OIG Comments

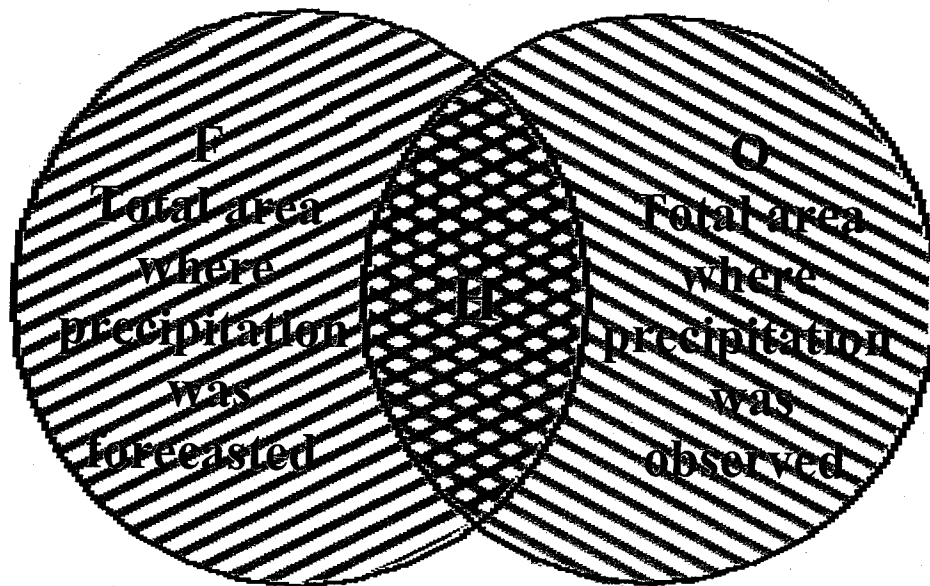
We commend NOAA and the NWS for the corrective actions taken and planned. We believe these actions demonstrate a commitment to the reporting of reliable performance data. Where appropriate, we have modified the report to reflect NOAA's response. The modifications included: (1) clarifying NWS' initial decision not to implement Rule 2, and (2) providing additional discussion of planned automated quality control upgrades. With respect to the NWS comments regarding policies that have been put into place to address the issue about including flood data with flash flood data, we have included references to these actions within the report.

III. Accuracy (%) of Three-day Forecast of Precipitation

NWS' Hydrometeorological Prediction Center (HPC) routinely prepares and distributes forecasts of precipitation for the contiguous United States.

To gauge how well NWS performs this activity, HPC measures how often (in percent) the Weather Service accurately forecasts precipitation 3 days in advance of the anticipated event and sets an annual goal for improving its success rate from one year to the next. HPC calculates the accuracy rate by dividing the area (in square kilometers) where precipitation was forecasted and observed by the total area where precipitation was observed.

Figure 1: Accuracy (%) of 3-Day Forecast of Precipitation—Current and Revised Measures



H= Area where precipitation was forecasted and observed

Current Measure
= H / O

Revised Measure—the Threat Score
= H / (F + O - H)

For FY 2001, those areas were 8.079 million square kilometers (area forecasted and precipitation observed) divided by 41.77 million square kilometers (total area of observed precipitation). Consequently, NOAA reported for fiscal year 2001 an accuracy rate of 19 percent in the *FY 2001 APPR/FY 2003 APP*.¹³ Its target for the year was 22 percent.

¹³ NOAA and the Department also reported this measure in the *FY 2002 Performance & Accountability Report*. NOAA explains in the *FY 2001 APPR/FY2003 APP* that drought conditions are highly correlated to lower performance scores for precipitation.

To assess the accuracy of the reported results, we selected a judgmental sample of HPC's daily quantitative precipitation logs for June, July, August, and September of 2001. We also compared the totals for each month with those used by HPC to calculate the measure. We assessed the usefulness of the measure and consideration of areas in which precipitation was forecasted but not observed.

A. Improvements are needed to more accurately measure the forecasting performance.

This measure is of limited usefulness because it does not take into account areas where rain is forecast but does not occur (see Figure 1), but only considers areas where precipitation occurs (whether forecasted or not). Therefore, it does not fully reflect NWS' ability to forecast precipitation 3 days in advance.

HPC management informed us that it had requested replacement of the current measure with a more useful one, called the "threat score," which takes into account areas where rain was correctly forecasted, incorrectly forecasted, or not forecast at all but did occur. NWS has indicated that it decided to replace the measure in December 2001.



Measuring rainfall
Source: NOAA Photo Library

A NWS official confirmed that the new measure was approved in January 2003 by the Under Secretary of Commerce for Oceans and Atmosphere and NOAA Administrator. We concur with NWS' action to revise this measure, and note that NOAA must explain the change when the new measure is used in future performance reports.

B. Additional disclosures are needed to improve presentation.

NOAA needs to provide additional explanations in the following areas to make the discussion of this measure more useful:

Data limitation. The discussion of data limitation in the *FY 2001 APPR/FY 2003 APP* does not disclose that (1) the results do not include the area where rain was forecasted but did not occur, and (2) the forecasts counted in the measure are for 1 inch or more of precipitation.

Data verification. The *FY 2001 APPR/FY 2003 APP* states that "all data are examined for accuracy, and quality control procedures are applied," but provides no specifics describing the procedures HPC uses to verify accuracy—such as checking for errors and eliminating duplicates. Adequate explanation of quality control procedures helps readers determine how reliable the data is.

C. Recommendations

The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA in subsequent reports

- (1) notes that the FY 2001 results did not reflect areas where precipitation was forecasted but did not occur;
- (2) states that the measure gauges forecasts of 1 inch or more; and
- (3) provides specific examples of data verification procedures.

D. NOAA Response

In response to the draft report, NOAA concurred with all three recommendations, identifying corrective actions taken or planned and target completions dates. NWS noted that it had decided to replace this measure in December 2001 with a more useful measure entitled "threat score." This new measure will be reported in the NOAA FY 2003 APPR. Subsequent reports will indicate that this measure gauges forecasts of one inch or more of precipitation and detail specific verification procedures for this measure.

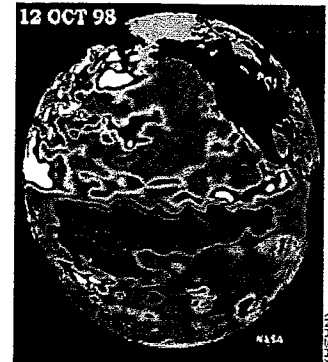
E. OIG Comments

We are encouraged by the NWS actions taken and planned with respect to this measure.

IV. Determine the Accuracy of the Correlation Between Forecasts of the Southern Oscillation Index (SOI) and El Niño/La Niña Events ¹⁴

Because ocean temperatures and circulation patterns influence the atmosphere, NOAA gathers sea surface temperature from buoys, ships, and satellites to help formulate its seasonal and interannual forecasts. Southern oscillation is an atmospheric phenomenon that greatly influences this process.

In both the *FY 2001 APPR/FY 2003 APP* and the *FY 2002 Performance & Accountability Report*, NOAA purports to assess the accuracy of the correlation between forecasts of South Pacific sea surface temperature and actual sea surface temperature connecting Tahiti and Darwin. The measure consists of 16 years of cumulative forecast data and observed actual sea surface temperatures. However, the measure does not demonstrate accuracy, but rather proximity—that is, how closely NWS forecasts correlate with observed temperatures.



In October 1998, cold La Nina water spans most of the equatorial Pacific following the strong El Nino of the previous years. Source: NASA

For FY 2001, NOAA reported meeting its target correlation index of .85. We independently tested the correlation between the forecast and actual sea surface temperatures for the 16 years beginning October 1985 and ending September 2001, and generated a correlation index of .85, confirming NOAA's reported results.

A. Usefulness of measure is hampered by limited, inadequate discussion of results.

NOAA does not clearly articulate what is being measured and does not provide enough pertinent discussion to make the reported information useful and meaningful. The presentation of FY 2001 data in the *APPR/APP* does not disclose that the measure is cumulative and includes 16 years of data, and does not explain the correlation index (.85). Understanding this index is essential to correctly interpreting the measure and putting NWS' performance in context. Without adequate explanation, the data's accuracy and reliability is open to question.

The title of the measure is incorrect as well. The data collected represents a relationship, not the accuracy of forecasts. Officials at NWS' National Centers for Environmental Prediction (which stores the data) stated that a more appropriate title would be "Determine the correlation of the forecasts of the Southern Oscillation Index (SOI) and El Niño/La Niña Events."

¹⁴ The surface air pressure differences observed between Tahiti and Darwin, Australia, known as the Southern Oscillation Index (SOI), are strongly linked to El Niño and La Niña. The SOI is frequently used as a convenient, simple, and reasonably accurate tool to monitor the status of El Niño and La Niña. El Niño is the warm phase of the Southern Oscillation, and La Niña is the cold phase of the Southern Oscillation.

In response to our concerns, NOAA enhanced the measure's explanation in the *FY 2002 Performance & Accountability Report*, stating that, "For the measure on correlation accuracy, the FY 2002 correlation was again computed using the past fifteen years of monthly values of forecast." Further, officials at the National Centers for Environmental Prediction are receptive to eliminating this measure from future annual performance reports, but retain it for internal reporting purposes.

B. Recommendations

The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA takes the following actions:

- (1) Remove the performance measure or revise the title and include appropriate discussion in future performance reports.
- (2) If the measure is revised, include necessary disclosures and explanation of changes in the presentation in future APPRs.

C. NOAA Response

In response to the draft report, NWS concurs with the two recommendations. NWS will amend the title of the measure and expand the explanation and disclosures for the measure in the FY 2003 APPR. Also, new performance measures are being evaluated for the NOAA climate program.

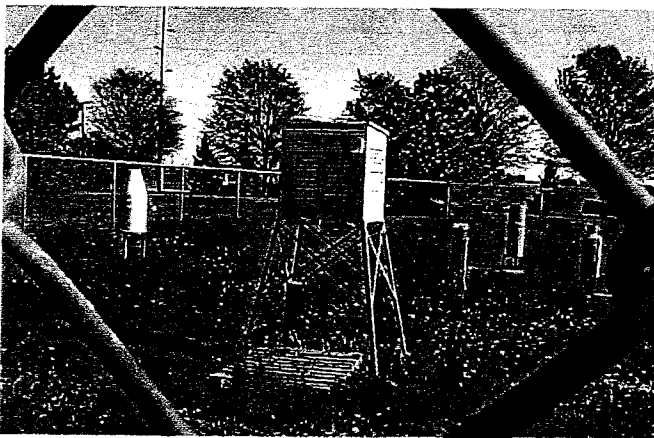
D. OIG Comments

We commend NWS for the actions it is taking with respect to this performance measure.

V. U.S. Temperature Forecasts (Skill Score)

This performance measure gauges NWS' skill at predicting temperature over the past 48 months against the random chance of being correct.

NOAA's target score since FY 1999 has been 20 (on a scale of -50 to +100),¹⁵ as reported in the *FY 2001 APPR/FY 2003 APP* and *FY 2001 Accountability Report*. NOAA reported surpassing its target in fiscal years 1999 and 2000, reporting scores of 23 and 25, respectively. For FY 2001, NOAA reported a skill score of 20 and thus met its target.



NWS temperature gauge.
Source: NOAA

We assessed the accuracy of the reported results by looking at five seasons from the period beginning July 2000 and ending September 2001, and comparing the skill scores in our sample with those NOAA used to calculate the cumulative average over 48 months. Also, we compared the actual skill scores for FY 1999 and FY 2000 with the reported results in the *FY 2001 APPR/FY 2003 APP* to determine whether there was consistency in reporting.

A. Performance results were calculated using different time spans.

Skill scores for FY 1999 and 2000 were calculated using a 36-month average—NWS' standard prior to FY 2001, when it was changed to 48 months. However, nowhere in either the *FY 2001 APPR/FY 2003 APP* or the FY 1999, FY 2000, and FY 2001 *Accountability Reports* is there mention of the monthly average used in the calculation, or that the average has been changed. Because reporting periods are not comparable, readers cannot analyze the data for trends, and the usefulness of the results is therefore limited.

When the actual FY 1999 and FY 2000 scores are recalculated using the 48-month average, the numbers reported for those years are misstated: in FY 1999 by 15 percent (the score was 19.88 rather than 23) and in FY 2000 by 7 percent (the score was 27 rather than 25).

¹⁵ NOAA uses the Heidke skill score to calculate this measure—considered the standard for forecasting by the scientific community. Based on Heidke's scale of -50 to +100, when forecasters match a random prediction the score is zero. Anything above zero shows positive skill in forecasting. Given the difficulty of forecasting temperatures in advance, a skill score of 20 is considered quite good.

B. Additional disclosures are needed to improve presentation.

NOAA's presentation of the measure in the *APPR/APP* does not provide sufficient discussion to give the reader a clear understanding of the reported results:

- NOAA does not reveal that the score covers forecasts for only about 40 percent of the U.S. The measure's title implies that the score covers the entire nation.
- NOAA does not explain the decline in skill score from FY 2000 to FY 2001—detail that would likely be of interest to decision makers and the public. However, NOAA took steps to address this concern in the Department's *FY 2002 Performance & Accountability Report*, noting that "NOAA Weather Service missed the target for the year—skill of seasonal prediction is influenced by the strength of predictors, El Niño being one. The El Niño pattern experienced in FY 2002 was weak-to-moderate, resulting in reduced overall accuracy of climate forecasts for the year."
- The results are calculated manually and there are no written procedures for verifying the calculations. Such procedures would help prevent the reporting of inconsistent or inaccurate performance data. Climate Prediction Center officials stated that written verification procedures could be added to the *Station Duty Manual*, which contains internal procedures used by meteorologists for developing their operational products.

Finally, as mentioned earlier, NOAA did not disclose in the *APPR/APP* that the skill score is a cumulative average of the past 4 years. It has since responded to this concern by noting in the Department's *FY 2002 Performance & Accountability Report* that "[t]he end of the year actual [score] represents a running average of mean score for the previous forty-eight months."

C. Recommendations

The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA takes the following actions:

- (1) Revise reported results for FY 1999 and FY 2000 and provide an explanation of the change from 36 to 48 months when reporting results for this measure in subsequent APPRs.
- (2) Disclose in APPRs that the measure is a cumulative average of 48 seasons and does not cover the entire U.S.
- (3) Develop and enforce procedures for ensuring that performance data is reviewed for accuracy prior to its inclusion in the APPR.

D. NOAA Response

In response to the draft report, NWS concurred with all the recommendations. NWS stated that (1) FY 1999 and FY 2000 results have been recomputed based on a 48-month cumulative average and will be provided in subsequent reports; (2) subsequent reports containing performance information will contain better explanations of the measure; and (3) the NWS Climate Prediction Center has already taken action to ensure the accuracy of the data, including improved quality control procedures and additional review of data prior to inclusion in subsequent reports. NWS states that its plans to improve its explanation of the measure meets the intent of the draft report recommendation to revise the title of the performance measure or consider use of a new measure.

E. OIG Comments

We commend NWS for the actions it is taking or plans to take with respect to this performance measure. We agree that the NWS actions are consistent with the intent of the recommendations and subsequently dropped the draft report recommendation to revise the title of the measure or create a new measure.

VI. Number of New Monitoring or Forecast Products that Become Operational Per Year (cumulative)

This performance measure reports new products for monitoring weather and development of new forecasts. NOAA reported in the *FY 2001 APPR/FY 2003 APP* that it issued four new products—its target number—in fiscal year 2001: (1) the Numerical Model Forecast Evaluation Product, used to develop 5-, 10-, and 15-day forecasts; (2) a product that monitors the Arctic Oscillation phenomena; (3) the heat index forecast product; and (4) a new wind chill forecast product.



		Temperature (°F)																		
		40	35	30	25	20	15	10	5	0	-5	-10	-15	-20	-25	-30	-35	-40	-45	-50
Wind (mph)	5	36	31	25	19	13	7	1	-5	-11	-16	-22	-28	-34	-40	-46	-52	-57	-63	-69
	10	34	27	21	15	9	3	-4	-10	-16	-22	-28	-35	-41	-47	-53	-59	-66	-72	-78
	15	32	25	19	13	6	0	-7	-13	-19	-26	-32	-39	-45	-51	-58	-64	-71	-77	-83
	20	30	24	17	11	4	-2	-9	-15	-22	-29	-35	-42	-48	-55	-61	-68	-74	-81	-87
	25	28	23	16	9	3	-4	-11	-17	-24	-31	-37	-44	-51	-58	-64	-71	-78	-84	-91
	30	26	22	15	8	1	-5	-12	-19	-26	-33	-39	-46	-53	-60	-67	-73	-80	-87	-94
	35	24	21	14	7	0	-7	-14	-21	-27	-34	-41	-48	-55	-62	-69	-76	-82	-89	-96
	40	22	20	13	6	-1	-8	-15	-22	-29	-36	-43	-50	-57	-64	-71	-78	-84	-91	-98
	45	20	19	12	5	-2	-9	-16	-23	-30	-37	-44	-51	-58	-65	-72	-79	-86	-93	-100
	50	18	19	12	4	-3	-10	-17	-24	-31	-38	-45	-52	-60	-67	-74	-81	-88	-95	-102
	55	16	18	11	3	-4	-11	-18	-25	-32	-39	-46	-54	-61	-68	-75	-82	-89	-96	-103
	60	14	17	10	2	-5	-12	-19	-26	-33	-40	-48	-55	-62	-69	-76	-84	-91	-98	-105

Frostbite Times ☐ 30 minutes ☐ 10 minutes ☐ 5 minutes

$$\text{Wind Chill (°F)} = 35.74 + 0.6215T - 35.75(V^{0.16}) + 0.4275T(V^{0.16})$$

Where: T = Air Temperature (°F) V = Wind Speed (mph)

Effective 11/01/02

NWS Wind Chill Temperature Index

Source: NWS

We interviewed the meteorologists from NWS' Climate Prediction Center (CPC) who developed the monitoring/forecast products that became operational in FY 2001, and examined what procedures—if any—were in place for verifying the accuracy of the results NOAA reported. We learned that depending on the definition used for a new product, the actual number of new products that became operational in FY 2001 could have been significantly understated in the *APPR/APP*.

A. NOAA needs a consistent definition for a monitoring or forecast product

Our audit found that there was no single, consistent definition for a new product at CPC. Depending on the definition used, 50 new products could have been identified as becoming operational in FY 2001—not 4, as reported in the *FY 2001 APPR/FY 2003 APP*. To ensure accurate reporting, NOAA needs to develop and disseminate such a single, consistent definition of “new product.” NWS indicated that it will implement an improved definition of a new product and that this will eliminate the confusion over what should be considered a new product.

B. NOAA has no verification procedures for new products.

NOAA has no formal procedures for verifying the new product numbers submitted by CPC and the National Climatic Data Center (NCDC). NOAA simply notes in the *APPR/APP* that “products are reported to NOAA management at quarterly reviews,” and, in the *FY 2001 Accountability Report*, that “[p]roducts are reported and reviewed on a quarterly basis.”

Clear, consistently applied verification procedures ensure that only valid and useful performance data is reported and that the data is reliable. CPC officials stated that written verification procedures could be added to the *Station Duty Manual*.

C. Recommendations

The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA develops procedures to verify the number of new monitoring/forecast products that become operational annually before reporting the results.

D. NOAA Response

In response to the draft report, NWS and NESDIS concurred with the recommendation that procedures be developed to verify the number of new monitoring/forecast products that become operational annually before reporting results. However, the NWS did not concur with the recommendation to revise FY 2001 results and adjust performance targets for the FY 2003 *APPR* as the NWS maintains that only four new products were developed during FY 2001. However, it recognizes a lack of a common definition for a new product at CPC resulted in confusion as to the number of products to be reported. The National Climatic Data Center (NCDC) will work with CPC to develop and disseminate a definition of "operational monitoring/forecast product" and will incorporate that definition in future reporting of new operational monitoring/forecast products.

E. OIG Comments

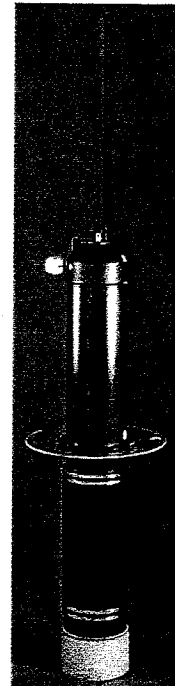
We commend NWS for its plans to develop and disseminate an improved definition for a new product for this performance measure. The planned efforts of the NWS to eliminate the confusion as to the definition of a new product, are consistent with the intent of our recommendations. As such, we have eliminated the recommendation for NWS to revise FY 2001 results and adjust performance targets for the FY 2003 *APPR*.

VII. New Climate Observations Introduced

The Office of Oceanic and Atmospheric Research and NESDIS are expanding their climate observation systems as part of NOAA's efforts to improve its climate monitoring and prediction capability. This measure is intended to record the number of new monitoring systems deployed or made operational during the fiscal year.

We examined NOAA's FY 2001 results, which report on deployment of the Argo network—a global array of free-drifting floats that measure temperature and salinity of the upper 2,000 meters of the ocean. Deployment of the floats began in FY 2001, and will ultimately number 3,000.

NOAA's FY 2001 target for this measure was 120 floats, and it reported introducing 132, thereby exceeding its goal. However, we found this number was inaccurate: it identifies the number of Argo floats budgeted for procurement, not deployed. In addition, the discussion of the measure refers to other observation systems as well (i.e., buoys, ships, and satellites). In response to these and other concerns we raised, NOAA began exploring alternative presentations of this measure prior to the conclusion of our audit. We encourage this effort and believe our findings here can help direct the search.

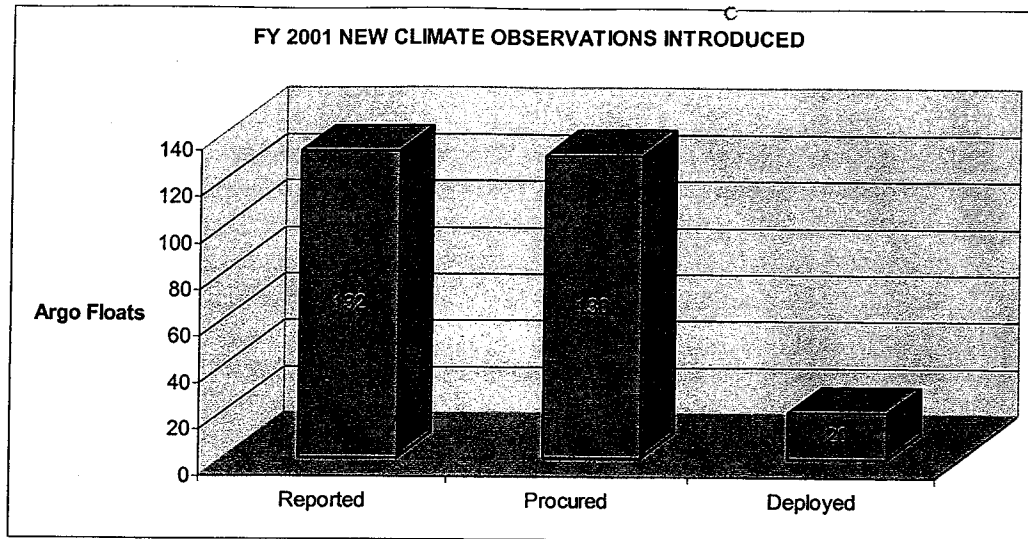


Argo float.
Source: Ocean US

A. Reported results are inaccurate

The 132 Argo floats reported as deployed in FY 2001 actually represent the number budgeted in FY 2000. Actual procurement occurred in FY 2001; the total purchased was 130 (owing to price increases). NOAA deployed 20 of the 130 in the year of purchase, and deployed the remainder in fiscal years 2002 and 2003.

Because NOAA does not explain in the *FY 2001 APPR/FY 2003 APP* that it is reporting the number of floats budgeted for procurement rather than deployed, the reader is led to believe that 132 floats have been newly put in place or made operational, and that the process is therefore well ahead of where it actually is. No one verified the accuracy of the performance results prior to its publication in the *APPR/APP*, and NOAA could not readily produce documentation to support the number of floats deployed. In discussing verification in the *APPR/APP*, NOAA states that it "performs quality assurance analysis," but this procedure is applied only to the data coming from the floats—the results are not verified, although the reader is led to believe otherwise. To ensure the quality, integrity, and validity of reported data, NOAA needs to be clear about the quality control procedures it performs, implement stronger quality controls, and have supporting documentation readily available.



B. Measure does not accurately convey what is being reported.

The title of the measure and its accompanying narrative imply that NOAA is counting multiple climate observation systems or products, and specifically mentions data buoys and new satellites, when in reality the reported number is measuring the introduction of Argo floats only. NOAA had other weather monitoring products deployed during FY 2001, including nine “Argo Equivalents”¹⁶ and weather platforms for the climate reference network. In FY 2002, NOAA deployed an even greater variety of new observing equipment including ocean reference moorings and carbon flux monitoring sites, but does not include them in the results reported for this measure, and therefore limits its usefulness. If NOAA plans to report only on Argo float deployments, it should revise the measure and discussion accordingly. If not, NOAA should include all new observation equipment deployed.

C. Recommendations

The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA takes the following actions:

- (1) Reevaluates the usefulness of this performance measure, and revises the discussion of results in future performance reports to include appropriate disclosures that clarify and enhance meaning.
- (2) Establishes procedures for reporting only appropriate fiscal year results, developing and maintaining adequate support documentation, and reconciling performance data with documentary evidence.

¹⁶ Argo-equivalent floats are essentially Argo floats that have been deployed and/or programmed differently.

D. NOAA Response

In its response to the draft audit report, OAR concurred with both recommendations. In future reports containing performance information, this measure will be adjusted to reflect only deployed ocean observing instruments and platforms. Also, OAR will (1) establish procedures for reporting only appropriate fiscal year results, (2) develop and maintain adequate supporting documentation, and (3) reconcile performance data with documentary evidence.

E. OIG Comments

We are encouraged by the OAR response as it indicates a commitment on the part of OAR to improve the reporting of this performance measure.



APPENDIX I

UNITED STATES DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
NOAA FINANCE AND ADMINISTRATION
CHIEF FINANCIAL OFFICER/CHIEF ADMINISTRATIVE OFFICER

JUL 14 2003

MEMORANDUM FOR: Johnnie E. Frazier
Inspector General

FROM: Helen Hurcombe *Helen Hurcombe*
Acting Chief Financial Officer/
Chief Administrative Officer

SUBJECT: *Improvements Needed in the Reporting of Performance
Measures Related to Goals for Advanced Short-Term
Warnings and Implementing Seasonal to Internannual
Climate Forecasts*
Draft Audit Report No. FSD-15643-3-0001/June 2003

Attached is the National Oceanic and Atmospheric Administration's response to the Office of Inspector General's draft audit report on the selected performance measures at the National Weather Service. The response has been prepared in accordance with Department Administrative Order 213-3.

We appreciate the opportunity to respond to your draft audit report.

Attachment



National Weather Service (NWS) Response to the
Office of the Inspector General (OIG) Draft Report
No. FSD-15643-3-0001, June 2003

*Improvements Needed in the Reporting of Performance Measures
Related to Goals for Advancing Short-Term Warnings and
Implementing Seasonal to Interannual Climate Forecasts*

Observations and Conclusions

Overall Comments

Overall, the tone and tenor of the draft OIG report is too negative compared to the relatively minor significance of the findings and recommendations. In most cases, NOAA was already aware of the issue and had implemented corrections prior to or during the IG Audit.

The IG report should include more information and details on the extensive processes used by NOAA to ensure the accuracy of performance information. For example, the NWS conducts site visits and maintains extensive documentation to verify the accuracy of weather warnings. In addition, the NWS maintains a centralized tracking database for performance information with automated quality controls. This database was established in response to an previous IG Review. The IG Report should reference some of these important operational policies and procedures to give the reader a more balanced view.

NOAA is also concerned about the sample size used to examine performance data for the tornado and flash flood lead times. For example, the IG report notes a recording error in 1 out of 20 tornado warnings and concluded the error was an internal control weakness and could impact the accuracy of National performance information. While every system is subject to human error, the IG sampling method was not statistically significant and may overstate the extent of the recording errors.

Response to OIG Draft Report No. FSD-15643-3-0001/June 2003

Finding I - Performance Measure: Lead time (minutes), accuracy (%), and false alarm rate (FAR, %) for severe weather warnings for tornadoes.

Recommendations - The Under Secretary of Commerce for Oceans and Atmosphere should ensure NOAA does the following:

1. Provides performance data to reflect the percentage of events in which the public is not provided a warning in time to take precautionary actions.

NWS Response: NWS concurs. NWS will develop a new performance measure to augment our current set of measures. The new measure, in concert with existing measures, should provide a clearer assessment on performance without focusing solely on zero lead times. NWS will also expand disclosure information in future Annual Program Performance Reports (APPR) and Annual Performance Plans (APP) regarding tornado lead times.

Target Date of Completion: NOAA's FY 2005 APP (October 2003)

2. Strengthens internal controls over performance reporting, to include a policy for maintaining accurate supporting documentation and procedures for reconciling data prior to reporting.

NWS Response: NWS and NOAA concurs.

- a. Section I B, paragraph 1, states a performance result was incorrectly recorded by NOAA headquarters prior to submission to the Department in the FY 2002 Performance and Accountability Report. NOAA Headquarters will implement quality control and review improvements to prevent future recording errors.

Section I B, paragraph 2, notes one instance where an event time for a tornado reported by a WFO was incorrectly entered into the database. While the sample size for this finding is very limited and does not provide a representative sample, NWS will add language in NWS Instruction 10-1605, Storm Data Preparation, to reemphasize the need for accurate entries in Storm Data. NWS will also develop a process to spot-check storm data entries for accuracy.

- b. Section I C, regarding document retention, the NWS will modify NWS Instruction 10-1605, Storm Data, to include a requirement for retaining all documentation used for the production of Storm Data for two years.

Target Date of Completion:

- a. September 30, 2003, or three months from release of OIG Final Report
- b. September 30, 2003, or three months from release of OIG Final Report

3. Revises the presentation of the measure to include all appropriate disclosures to make the discussion of results more clear and meaningful.

NWS Response: NWS concurs. Section I D, states NWS needs to establish written guidelines to fully disclose source and quality of data. NWS agrees and has already modified the description for this measure in the draft NOAA FY 2005 APP. This information will continue to be reported in future submissions.

Target Date of Completion: NOAA's FY 2005 APP (October 2003)

4. Takes other appropriate actions--such as upgrading systems and software--to facilitate and ensure the reporting of accurate data.

NWS Response: NWS concurs.

- a. NWS has already taken significant steps to facilitate and ensure the accuracy of reporting data. As noted in the OIG report, NWS has implemented automated algorithms, entitled Rule 1 and Rule 2, to ensure accuracy of performance information and provide WFOs with real-time feedback when errors are detected in warnings. Pages ii, iii, and 9 make reference to a temporary suspension of quality control procedures for the FY 2002 data. However, in the transition from manual to automated quality control, Rule 1, an automated procedure to check for warning coding and format errors, was in place (see Appendix). We agree Rule 2, an automated procedure to quality control overlapping and corrected warnings, was implemented before the final FY 2002 numbers were calculated, ensuring the accuracy of data. We recommend the OIG report reference the implementation of Rule 1 in January of 2002 and qualify the wording regarding the suspension of control checks.
- b. New software is scheduled for inclusion in AWIPS Operational Build 2, scheduled to be released in December 2003. The software will include a Valid Time Event Code (VTEC) capability which will allow tracking of warnings by event number and identification of every warning as new or corrected. Quality control checks have also been added to the software that generates warnings. These checks include safeguards to ensure proper coding and formatting of warnings.

Target Date of Completion:

- a. Completed. Rule 1 implemented January 2002; Rule 2 implemented December 2002.
- b. AWIPS Operational Build 2 in December of 2003.

Finding II - Performance Measure: Lead time (minutes) and accuracy (%) for severe weather warnings for flash floods.

Recommendations - The Under Secretary of Commerce for Oceans and Atmosphere should ensure NOAA does the following:

- 1. Provide performance data to reflect the percentage of events in which the public is not provided a warning in time to take precautionary actions.

NWS Response: NWS concurs. NWS will develop a new performance measure to augment our current set of measures. The new measure, in concert with existing measures, should provide a clearer assessment on performance without focusing solely on zero lead times. NWS will also expand disclosure information in future Annual Program Performance Reports (APPR) and Annual Performance Plans (APP) regarding flash flood lead times.

Target Date of Completion: NOAA's FY 2005 APP (October 2003)

- 2. Strengthen internal controls for the performance measure, to include a policy for maintaining accurate supporting documentation and procedures for reconciling data prior to reporting.

NWS Response: NWS concurs.

- a. Section II B, bullet 1, regarding mixing of flash flood and flood data, NWS has already issued policies to address this issue. The policy prohibiting issuance of flash flood warnings for general area flooding was implemented on August 6, 2001. Definitions for flash floods and floods are contained in NWS Instruction 10-950, September 26, 2002. Policy to include flash flood warnings only and flash flood events only in the NWS flash flood verification program was established with issuance of NWS Instruction 10-1605, January 6, 2003.
- b. Section II B, bullet 2, incorrect event times: NWS will add language in NWS Instruction 10-1605, Storm Data Preparation, to reemphasize the need for accurate entries in Storm Data. NWS will develop process to spot-check storm data entries for accuracy.

- c. Section II C, supporting document retention: NWS will modify NWS Instruction 10-1605, Storm Data, to include a requirement for retaining all documentation used for the production of Storm Data for two years.

Target dates of completion:

- a. Completed January 6, 2003
- b. September 30, 2003, or three months from release of OIG Final Report
- c. September 30, 2003, or three months from release of OIG Final Report

3. Revise the presentation of the measure to include all appropriate disclosures to make the discussion of results more clear and meaningful.

NWS Response: NWS concurs. Regarding Section II D, establish written guidelines to fully disclose source and quality of data, NWS has modified descriptions in our portion of NOAA's FY 2005 APP to disclose what the data represents. This information will continue to be reported in future submissions.

Target Date of Completion: NOAA's FY 2005 APP (October 2003)

4. Take other appropriate actions--such as upgrading systems and software--to facilitate and ensure the reporting of accurate data.

NWS Response: NWS concurs. Section II B, bullet 4, warning mistakenly issued:

- a. NWS has already taken significant steps to facilitate and ensure the reporting of accurate data. As noted in the OIG report, NWS has implemented automated algorithms, entitled Rule 1 and Rule 2, to ensure accuracy of performance information and provide WFOs with real-time feedback when errors are detected in warnings. Pages ii, iii, and 9 make reference to a temporary suspension of quality control procedures for the FY 2002 data. However, in the transition from manual to automated quality control, Rule 1, an automated procedure to check for warning coding and format errors, was in place (see Appendix). We agree Rule 2, an automated procedure to quality control overlapping and corrected warnings, was implemented before the final FY 2002 numbers were calculated, ensuring the accuracy of data. We recommend the OIG report reference the implementation of Rule 1 in January of 2002 and qualify the wording regarding the suspension of control checks.

- b. New software is scheduled for inclusion in AWIPS Operational Build 2 to be released in December 2003. It includes a Valid Time Event Code (VTEC) capability which will allow tracking of warnings by event number and identification of every warning as new or corrected. Quality control checks have been added to the software that generates warnings. These checks include safeguards to ensure proper coding and formatting of warnings.

Target Date of Completion:

- a. Completed. Rule 1 implemented January 2002; Rule 2 implemented December 2002.
- b. AWIPS Operational Build 2 in December 2003.

Finding III - Accuracy (%) of three-day forecast of precipitation.

Recommendations - The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA in subsequent reports:

1. Note that the FY 2001 results did not reflect areas where precipitation was forecasted but did not occur.

NWS Response: NWS concurs. NWS had recognized the limitations of this measure and decided to replace the measure in December 2001. The change was reflected in the FY 2003 NWS Annual Operating Plan with a more useful measure entitled "threat score" which takes into account areas where one inch or more of precipitation was correctly forecasted, where it was forecasted but did not occur, and where it occurred but had not been forecasted. Subsequent reports will explain the change to this measure.

Target Date of Completion: NOAA's FY 2003 APPR (October 2003)

2. State that the measure gauges forecasts of one inch or more.

NWS Response: NWS concurs. Subsequent reports will indicate this measure gauges forecasts of one inch or more of precipitation

Target Date of Completion: NOAA's FY 2003 APPR (October 2003)

3. Provide specific examples of data verification procedures.

NWS Response: NWS concurs. Subsequent reports will contain specific data verification procedures for this measure.

Target Date of Completion: NOAA's FY 2003 APPR (October 2003)

Finding IV - Determine the accuracy of the correlation between forecasts of the Southern Oscillation Index (SOI) and El Niño/La Niña events.

Recommendations - The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA takes the following actions:

1. Remove the performance measure or revise the title and include appropriate discussion in future performance reports.

NWS Response: NWS concurs. NWS will amend the title and expand the explanation in the FY 2003 APPR. NOAA is also considering eliminating external reporting of this measure in the FY 2005 APP. The measure may be too technical for the broader NOAA climate audience. NOAA is evaluating new performance measures for the NOAA climate program.

Target Date of Completion: NOAA's FY 2003 APPR (October 2003)

2. If the measure is revised, include necessary disclosures and explanation of changes in the presentation in future APPRs.

NWS Response: NWS will expand the explanation and include a description of the changes in the FY 2003 APPR.

Target Date of Completion: NOAA's FY 2003 APPR (October 2003)

Finding V - U.S. temperature forecasts (skill score).

Recommendations - The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA takes the following actions:

1. Revise reported results for FY 1999 and FY 2000 and provide an explanation of the change from 36 to 48 months when reporting results for this measure in subsequent APPRs.

NWS Response: NWS concurs. FY 1999 and FY 2000 results have been recomputed based on a 48-month cumulative average and will be provided in subsequent reports.

Target Date of Completion: NOAA's FY 2003 APPR (October 2003)

2. Disclose in APPRs that the measure is a cumulative average of 48 seasons and does not cover the entire U.S.

NWS Response: NWS concurs. Areas where no forecast for surface temperature is made (i.e., areas designated as "equal chance" on the Climate Prediction Center seasonal forecast maps) are not included in the computations for this measure. Subsequent reports will contain better explanation of this measure. By doing this, it meets the intent of Recommendation V-3 below.

Target Date of Completion: NOAA's FY 2003 APPR (October 2003)

3. Revise the title of the performance measure to Selected U.S. Temperature Forecasts (Skill Score), or consider using a new performance measure that will be more meaningful and easier to understand.

NWS Response: NWS concurs. See response to Recommendation V-2 above.

Target Date of Completion: NOAA's FY 2003 APPR (October 2003).

4. Develop and enforce procedures for ensuring that performance data is reviewed for accuracy prior to its inclusion in the APPR.

NWS Response: NWS concurs. The NWS Climate Prediction Center has already implemented improvements to ensure the accuracy of the performance data, including improved quality control procedures and additional personnel reviewing the data prior to inclusion in subsequent reports.

Target Date of Completion: Completed February 2003

Finding VI - Number of new monitoring or forecast products that become operational per year (cumulative).

Recommendations - The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA:

1. Develops procedures to verify the number of new monitoring/forecast products that become operational annually before reporting the results.

NWS and NESDIS Response: NWS and NESDIS concur. NOAA will develop and implement a consistent procedure to better define and quality control the number of new climate products. Of note, NESDIS has procedures to address this issue. When a new product is produced by the National Climatic Data Center (NCDC) and made available to the public, on-line and/or off-line, NCDC assigns a Federal Geographic Data Committee (FGDC) number to the product and adds the product to its product catalog. The product is then accessible through the NOAA National Data Center Online Store and via NCDC's Web site. NCDC also maintains a record of new products by fiscal year with the associated FGDC number for audit verification purposes. NCDC will work with the Climate Prediction Center (CPC) to develop and disseminate a definition of "operational monitoring/forecast product" and will incorporate that definition in future reporting of new operational monitoring/forecast products.

Target Date of Completion: September 2003

2. Revises FY 2001 results and adjust performance targets for the FY 2003 APPR.

NWS Response: NWS does not concur. Discrepancy in the number of new products reported by NOAA and revealed during the audit was due to the lack of a common definition for new products at CPC. NWS maintains only four major new products were developed during FY 2001. The counting of separate components within the new products as unique new products led to a dramatic increase in perceived number of new products during the audit. Once NWS has implemented an improved definition for a new product, this will eliminate the confusion.

Target Date of Completion: September 2003

Finding VII - New climate observation introduced.

Recommendations - The Under Secretary of Commerce for Oceans and Atmosphere should ensure that NOAA takes the following actions:

1. Reevaluates the usefulness of this performance measure, and revises the discussion of results in future performance reports to include appropriate disclosures that clarify and enhance meaning.

OAR Response: OAR concurs. In the FY 2003 APPR/FY 2005 APP OAR will:

- a. Adjust the performance measure to reflect all of the new, in situ ocean observing instruments and platforms (profiling floats, drifting buoys, Volunteer Observing Ships, etc.) supporting seasonal to interannual climate prediction actually being deployed, serviced, and/or equipped.
- b. Report deployed vs. budgeted performance information and expand on the explanation of this measure.

Target Date of Completion:

- a. NOAA's FY 2003 APPR/FY 2005 APP (October 2003)
- b. NOAA's FY 2003 APPR/FY 2005 APP (October 2003)

2. Establishes procedures for reporting only appropriate fiscal year results, developing and maintaining adequate support documentation, and reconciling performance data with documentary evidence.

OAR Response: OAR concurs. OAR will establish procedures for reporting only appropriate fiscal year results, develop and maintain adequate support documentation, and reconcile performance data with documentary evidence. Of note, this recommendation refers to the administrative monitoring and reporting of the performance measure and not the scientific and technical quality control processing of the data. The scientific and technical quality control procedures are

conducted within an exhaustive process implemented according to international standards that have been developed in conjunction with the operational centers and other users of the observations in the United States and throughout the world.

Target Date of Completion: NOAA's FY 2003 APR/FY 2005 APP
(October 2003)

Subject: [Fwd: [Fwd: Re: [Fwd: Revised Warning Document]]]

Date: Mon, 30 Jun 2003 13:21:01 -0400

From: Donald Wernly <Donald.Wernly@noaa.gov> Internal

Organization: DOC/NOAA/NWS - National Weather Service

To: Nicholas Scheller <Nicholas.Scheller@noaa.gov>,

Steven Gallagher <Steven.Gallagher@noaa.gov>,

Sheila Beehler <Sheila.Beehler@noaa.gov>

Folks

Here is the documentation for rule 1.

Don

----- Original Message -----

Subject: [Fwd: Re: [Fwd: Revised Warning Document]]

Date: Mon, 30 Jun 2003 09:51:28 -0400

From: William Lerner <William.Lerner@noaa.gov>

Organization: DOC/NOAA/NWS - National Weather Service

To: Donald Wernly <Donald.Wernly@noaa.gov>

Don,

This is the only thing I can find. I'll check with Robb to see if he has anything else.

Remember, there was never any disagreement about Rule 1 so there isn't much of a paper trail.

Bill

----- Original Message -----

Subject: Re: [Fwd: Revised Warning Document]

Date: Fri, 21 Dec 2001 13:25:24 -0500

From: "Brent Macaloney" <Brent.Macaloney@noaa.gov>

To: Richard Smith <Richard.Smith@noaa.gov>

CC: William Lerner <William.Lerner@noaa.gov>, Robb Kookaby <Robb.Kookaby@noaa.gov>

References: <3C221E0F.735E8583@noaa.gov>

Here ya go!

-Brent

William Lerner wrote:

> Brent,

> Can you do this?

>

> Richard Smith wrote:

>

> > Bill,

> >

> > Would it be possible to get a revised version of the PDF document you
> > sent with Rules 1 and 2? We would like to forward this to our offices,
> > but without Rule 2 included for now, and I do not have the capability to
> > edit the document.

> >

> > Thanks!

> >

> > Rick

OVERVIEW

The number of overlapping and erroneous severe weather and flash flood warnings has increased during the past few years. To try and reduce the impact on customers, and at the same time more accurately measure what we issue, the performance branch has developed a set of guidelines on how warnings will be archived. Examples are provided to show how the warnings will be entered into our database. Note however these are only a few examples of what we encounter daily.

RULE I. - HOW WARNINGS WILL BE ENTERED INTO THE DATABASE

All data imported into the warning database will be taken directly from the warning. Data will not be entered into the database from any other information listed in the header/text of the warning. See the examples below.

WUUS53 KWSH 010000
SVRWSH
DCC001-003-005-010200-

<----- WARNING TYPE and WFO
<----- COUNTY and STATE WARNED

BULLETIN - EAS ACTIVATION REQUESTED
SEVERE THUNDERSTORM WARNING
NATIONAL WEATHER SERVICE SILVER SPRING MD
700 PM EST WED JAN 1 2001

<----- DATE AND ISSUANCE TIME

THE NATIONAL WEATHER SERVICE IN SILVER SPRING HAS ISSUED A

* SEVERE THUNDERSTORM WARNING FOR...
WASHINGTON COUNTY IN THE DISTRICT OF COLUMBIA
JEFFERSON COUNTY IN THE DISTRICT OF COLUMBIA
REAGAN COUNTY IN THE DISTRICT OF COLUMBIA

* UNTIL 900 PM EST

<----- EXPIRATION TIME

* AT 700 PM CST...SILVER SPRING DOPPLER RADAR INDICATED A SEVERE
THUNDERSTORM 2 MILES WEST OF ADAMS MORGAN...MOVING EAST AT 15
MPH.

THE SEVERE THUNDERSTORM IS CAPABLE OF PRODUCING...
HAIL THE SIZE OF NICKELS

LAT...LON 3778 9752 3748 9752 3749 9724 3785 9724

Rule I. - Example 1

WUUS53 KWSH 010000
SVRWSH
DCC001-003-005-010200-

BULLETIN - EAS ACTIVATION REQUESTED
TORNADO WARNING
NATIONAL WEATHER SERVICE SILVER SPRING MD
700 PM EST WED JAN 1 2001

THE NATIONAL WEATHER SERVICE IN SILVER SPRING HAS ISSUED A

- * TORNADO WARNING FOR...
WASHINGTON COUNTY IN THE DISTRICT OF COLUMBIA
JEFFERSON COUNTY IN THE DISTRICT OF COLUMBIA
REAGAN COUNTY IN THE DISTRICT OF COLUMBIA
- * UNTIL 900 PM EST
- * AT 700 PM CST...SILVER SPRING DOPPLER RADAR INDICATED A TORNADO
2 MILES SOUTHWEST OF ADAMS MORGAN...MOVING EAST AT 15 MPH.

LAT...LON 3778 9752 3748 9752 3749 9724 3785 9724

Even though the text states that it is a tornado warning, this warning will be entered into the database as a Severe Thunderstorm Warning due to its being labeled an SVR in the product identification header. The product was issued as a Severe Thunderstorm Warning and will be databased as a Severe Thunderstorm Warning.

Rule I. - Example 2

WUUS53 KWSH 010000
TORWSH
DCC001-003-005-009-021-010200-

BULLETIN - EAS ACTIVATION REQUESTED
TORNADO WARNING
NATIONAL WEATHER SERVICE SILVER SPRING MD
700 PM EST WED JAN 1 2001

THE NATIONAL WEATHER SERVICE IN SILVER SPRING HAS ISSUED A

- * TORNADO WARNING FOR...
WASHINGTON COUNTY IN THE DISTRICT OF COLUMBIA
JEFFERSON COUNTY IN THE DISTRICT OF COLUMBIA
REAGAN COUNTY IN THE DISTRICT OF COLUMBIA
LINCOLN COUNTY IN THE DISTRICT OF COLUMBIA
- * UNTIL 900 PM EST
- * AT 700 PM CST...SILVER SPRING DOPPLER RADAR INDICATED A TORNADO
2 MILES SOUTHWEST OF ADAMS MORGAN...MOVING EAST AT 15 MPH.

LAT...LON 3778 9752 3748 9752 3749 9724 3785 9724

In this tornado warning, notice there are four counties listed in the body of the text and five counties listed in the UGC code. In this case, the automated system will enter five tornado warnings into the database, even though there are four counties listed in the body of the warning.

Rule I. - Example 3

WUUS53 KWSH 010000
SVRWSH
DCC-001-005-010200-

BULLETIN - EAS ACTIVATION REQUESTED
SEVERE THUNDERSTORM WARNING
NATIONAL WEATHER SERVICE SILVER SPRING MD
700 PM EST WED JAN 1 2001

THE NATIONAL WEATHER SERVICE IN SILVER SPRING HAS ISSUED A

* SEVERE THUNDERSTORM WARNING FOR...
WASHINGTON COUNTY IN THE DISTRICT OF COLUMBIA
REAGAN COUNTY IN THE DISTRICT OF COLUMBIA

* UNTIL 900 PM EST

* AT 700 PM CST...SILVER SPRING DOPPLER RADAR INDICATED A SEVERE
THUNDERSTORM 2 MILES WEST OF ADAMS MORGAN...MOVING EAST AT 5
MPH.

THE SEVERE THUNDERSTORM IS CAPABLE OF PRODUCING...
HAIL THE SIZE OF NICKELS

LAT...LON 3778 9752 3748 9752 3749 9724 3785 9724

Notice in the UGC header there is an extra dash before county FIPS. The UGC line should read DCC001-005-010200-. This is an invalid warning and will not be added to our warning database.

Rule I. - Example 4

WGUS56 KWSH 010000
FFWWSH
DCC001-005-010200-

BULLETIN - EAS ACTIVATION REQUESTED
FLASH FLOOD WARNING
NATIONAL WEATHER SERVICE SILVER SPRING MD
700 PM WED JAN 1 2001

THE NATIONAL WEATHER SERVICE IN SILVER SPRING HAS ISSUED A

* FLASH FLOOD WARNING FOR...
WASHINGTON COUNTY IN THE DISTRICT OF COLUMBIA
REAGAN COUNTY IN THE DISTRICT OF COLUMBIA

* UNTIL 900 PM EST

* AT 700 PM EST...WEATHER SERVICE DOPPLER RADAR INDICATED VERY
HEAVY RAIN OVER ADAMS MORGAN. THE AREA OF HEAVY RAIN WAS NEARLY
STATIONARY.

DO NOT DRIVE YOUR VEHICLE INTO AREAS WHERE THE WATER COVERS THE
ROADWAY. THE WATER DEPTH MAY BE TOO GREAT TO ALLOW YOUR CAR TO
CROSS SAFELY. VEHICLES CAUGHT IN RISING WATER SHOULD BE ABANDONED
QUICKLY. MOVE TO HIGHER GROUND.

LAT...LON 3321 11738 3267 11714 3289 11684 3319 11668

Notice in the "date and issuance time" line there is no time zone indicator. There is no way to tell if this warning was issued in EDT, CST, or any other time zone. This is an invalid warning and will not be added to our warning database.

Rule I. - Example 5

WUUS53 KWSH 010000
SVRWSH
DCC001-005-010200-

BULLETIN - EAS ACTIVATION REQUESTED
SEVERE THUNDERSTORM WARNING
NATIONAL WEATHER SERVICE SILVER SPRING MD
700 PM EST WED JAN 1 2001

THE NATIONAL WEATHER SERVICE IN SILVER SPRING HAS ISSUED A

* SEVERE THUNDERSTORM WARNING FOR...
WASHINGTON COUNTY IN THE DISTRICT OF COLUMBIA
REAGAN COUNTY IN THE DISTRICT OF COLUMBIA

* UNTIL 9005 PM EST

* AT 700 PM CST...SILVER SPRING DOPPLER RADAR INDICATED A SEVERE
THUNDERSTORM 2 MILES WEST OF ADAMS MORGAN...MOVING EAST AT 15
MPH.

THE SEVERE THUNDERSTORM IS CAPABLE OF PRODUCING...
HAIL THE SIZE OF NICKELS

LAT...LON 3778 9752 3748 9752 3749 9724 3785 9724

Notice in the "expiration time" line that the time is not valid. This is an invalid warning and will not be added to our warning database.

Rule I. - Example 6

WUUS53 KWSH 010000
SVRWSH

BULLETIN - EAS ACTIVATION REQUESTED
SEVERE THUNDERSTORM WARNING
NATIONAL WEATHER SERVICE SILVER SPRING MD
700 PM EST WED JAN 1 2001

THE NATIONAL WEATHER SERVICE IN SILVER SPRING HAS ISSUEDcA

* SEVERE THUNDERSTORM WARNING FOR...
WASHINGTON COUNTY IN THE DISTRICT OF COLUMBIA
REAGAN COUNTY IN THE DISTRICT OF COLUMBIA

* UNTIL 900 PM EST

* AT 700 PM CST...SILVER SPRING DOPPLER RADAR INDICATED A SEVERE
THUNDERSTORM 2 MILES WEST OF ADAMS MORGAN.

THE SEVERE THUNDERSTORM IS CAPABLE OF PRODUCING...
HAIL THE SIZE OF NICKELS

LAT...LON 3778 9752 3748 9752 3749 9724 3785 9724

In this example, the UGC code is missing. The automated warning databasing system can not identify the counties. This is an invalid warning and will not be added to our warning database.

Rule I. - Example 7

WGUS56 KWSH 010000
FFWWSH
DCC001-05-010200-

BULLETIN - EAS ACTIVATION REQUESTED
FLASH FLOOD WARNING
NATIONAL WEATHER SERVICE SILVER SPRING MD
700 PM EST WED JAN 1 2001

THE NATIONAL WEATHER SERVICE IN SILVER SPRING HAS ISSUED A

* FLASH FLOOD WARNING FOR...
WASHINGTON COUNTY IN THE DISTRICT OF COLUMBIA
REAGAN COUNTY IN THE DISTRICT OF COLUMBIA

* UNTIL 900 PM EST

* AT 700 PM EST...WEATHER SERVICE DOPPLER RADAR INDICATED VERY
HEAVY RAIN OVER ADAMS MORGAN. THE AREA OF HEAVY RAIN WAS NEARLY
STATIONARY.

DO NOT DRIVE YOUR VEHICLE INTO AREAS WHERE THE WATER COVERS THE
ROADWAY. THE WATER DEPTH MAY BE TOO GREAT TO ALLOW YOUR CAR TO
CROSS SAFELY. VEHICLES CAUGHT IN RISING WATER SHOULD BE ABANDONED
QUICKLY. MOVE TO HIGHER GROUND.

LAT...LON 3321 11738 3267 11714 3289 11684 3319 11668

Notice how the second county FIPS listed in the UGC line only contains two digits. All county FIPS must have three digits. This is an invalid warning and will not be added to our warning database.